The Practical use of Cosmic Shear as a Probe of Gravity

Donnacha Kirk

Thesis submitted for the Degree of Doctor of Philosophy
of the University College London

Department of Physics & Astronomy
UNIVERSITY COLLEGE LONDON

April 2011

I, Donnacha Kirk, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Specifically:

- The work on luminosity functions and red galaxy fractions in section 2.7 was done by Dr. Michael Schneider (then at Durham) as part of our collaboration on the paper Kirk *et al.* (2010).
- The work in chapters 3 & 4 is my contribution to, and will form the substantial part of, papers in preparation with collaborators Prof. Rachel Bean and Istvan Lazslo at Cornell University. The MG effect on the linear growth function was computed by Istvan Lazslo.
- The Great08 section of chapter 6 draws on my contributions to the Great08 Handbook (Bridle *et al.* 2009), Great08 Results Paper (Bridle *et al.* 2010*a*) and Great10 Handbook (Kitching *et al.* 2010).

To Mum and Dad, for everything.

Boyle. ... an' it blowed, an' blowed – blew is the right word, Joxer, but blowed is what the sailors use...

Joxer. Aw, it's a darlin' word, a daarlin' word.

Boyle. An', as it blowed an' blowed, I often looked up at the sky an' assed meself the question – what is the stars, what is the stars?

Voice of Coal Vendor. Any blocks, coal-blocks; blocks, coal-blocks!

Joxer. Ah, that's the question, that's the question – what is the stars?

Boyle. An' then, I'd have another look, an' I'd ass meself – what is the moon?

Joxer. Ah, that's the question – what is the moon, what is the moon?

— Juno and the Paycock

Act I

ABSTRACT

This thesis deals with the practical application of Weak Gravitational Lensing (WGL) as a tool for studying two of the most challenging problems in modern cosmology: the nature of dark energy and tests of Einstein's General Relativity. To be a useful tool, WGL must be able to put robust constraints on the parameters of interest even in the presence of a strong systematic effect like the Intrinsic Alignment (IA) of galaxies.

In chapter 2 we demonstrate how ignoring the effect of IAs will lead to strongly biased measurements of cosmology. We introduce a simple IA model and demonstrate how the bias can be removed, at the cost of weaker constraints, by marginalising over nuisance parameters in the model. We present the first joint constraints on IAs and cosmology from WGL and galaxy position information. The combination returns errorbars of the same order as the naive WGL calculation which ignored IAs.

We extend the IA model to treat shear and position correlations in a unified way with > 100 nuisance parameters. We parameterise deviations from GR and show, in chapter 3, how the considerable ($\sim$75%) decrease in constraining power caused by IAs can be partially mitigated by the inclusion of position-position and shear-position correlations. In chapter 4 we find these results to extend across a wide range of survey specifications. We find that increased survey area is not always favoured at the expense of source density and that the impact of improved redshift knowledge is not dramatic.

In chapters 5-7 we note the potentially very powerful constraints from a joint analysis of WGL and Redshift Space Distortions (RSDs) from a spectroscopic survey. We summarise some important preparatory work for future cosmic shear surveys and advances in shear measurement techniques before concluding and describing some directions for future work.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

21st century cosmology, with its large teams of international collaborators, expensive telescopes, satellites and missions that can take decades from planning to final data release, is coming increasingly to resemble that most "industrial" of 20th century disciplines, high energy particle physics. Even so, modern astronomers are recognisably working in a tradition which comprises arguably the most ancient science of all. Astronomy was practised in every ancient civilisation and was the only practical branch of natural philosophy to be included in the quadrivium of the medieval universities alongside arithmetic, geometry and music.

Cosmology itself is perhaps the oldest and most compelling impetus to the study of astronomy. Pre-modern cosmology was indivisible from metaphysical speculation about the ordering of creation. Cosmos, itself cognate to order and harmony, is the divine reworking of Milton's *waste wide anarchy of Chaos*[1]. We may not be justified in seeing parallels with the formation of structure from initial perturbations but the works from Hesiod through Dante to Milton speak of an unflagging interest in the coherent description of what we would call our Universe. Cosmology today has shrugged off its theological and mystic origins, as astronomy is divorced from astrology, but the continued popularity of cosmology in print or on screen proves that it continues to answer a desire to know about the physical reality we inhabit, "For every one, as I think, must see that astronomy compels the soul to look upwards and leads us from this world to another." [2]

There is of course a counter-tradition to this high-minded reverence of the cosmologist-as-sage. We can be sure that Sophocles' thought: "Astronomy? Impossible to understand and mad-

[1] Paradise Lost, Book X

[2] Plato's Republic, Book VII, CUP, 529a

ness to investigate," has been echoed by many physicists over the centuries since and, as astronomy was to physics, so cosmology was to astronomy.

Even in the 20th Century, when the theory of Quantum Electrodynamics was achieving an accuracy in predictions of one part in a trillion, cosmologists were often content to be within a few orders of magnitude of the right answer. In part this is due to the nature of our enterprise. Unable to recreate the vast scales, energies or distances we study in a lab, cosmologists' only current option is to peer out into the Universe with ever better telescopes, probing a range of wavelengths of visible radiation. The Universe we observe is the only one we know exists. This means that we are stuck with some of the contingent properties of this realisation of post-Big Bang spacetime- we cannot run the Universe again from its initial conditions and see what turns out differently. We call the statistical limits this places on our measurements, *cosmic variance*.

Nevertheless, the progress that cosmology made over the 20th Century was striking. In the year 1900, before the work of Einstein and Hubble was published, there was no understanding of the expanding Universe, no real concept of the scale of the Universe beyond our galaxy and no advance on Newton's theory of gravity. By the end of the century we had produced an extremely successful "concordance cosmology". This concordance cosmology would seem very strange to a cosmologist from a century ago. Measurements of the CMB confirm an expanding Universe in the aftermath of a hot Big Bang. In fact the Universe is not only expanding but it is accelerating under the influence of some dark energy. This fluid, about which we know little, along with a type of non-interacting dark matter together make up $\sim$95% of the contents of our Universe.

The success of the concordance cosmology was made possible, in large part, by the emergence of precision cosmology in the last decade of the 20th Century. For the first time large datasets for multiple cosmological probes gave cosmologists the statistical power to accurately constrain a handful of cosmological parameters. With the available data set to increase by orders of magnitude in the coming decade, cosmology will never again be a guessing game or purely speculative pastime.

The probes, which are ready to exploit the explosion in available data, include the Cosmic Microwave Background, Weak Gravitational Lensing and Galaxy Redshift Surveys. To fully exploit the potential of the new datasets, our understanding of errors and systematic effects will also have to undergo a step change but, when we do fully exploit the new data, we will be able to test our cosmological models to new levels of precision. One of the most exciting research areas will be stringent tests of our understanding of gravity on cosmic scales. Einstein's General Relativity is an extremely successful theory but inconsistencies with quantum mechanics and the Universe's

unusual expansion history suggest that there may be the potential to understand new physics by probing the nature of gravity in the wider Universe.

The work in this thesis has been conducted in the evolving context of early 21st Century cosmology. One of the most important developments in the field is that it is becoming increasingly rare for observational cosmologists to work solely independently or with a very small team. Much work, particularly ambitious survey missions, is now carried out by international teams of astronomers which often number hundreds of members from tens of institutions. The context of these collaborations, design studies, technical potential, trial simulations and eventually, data, shape a lot of the research priorities in the field (Lahav 2001).

This introduction aims to familiarise the reader with some of the basic concepts which underpin the work in the following chapters. 1.1 treats the modern cosmological model in general, including the smooth expanding Universe and the FRLW metric (1.1.1), our understanding of gravity (1.2), the dynamics of the Universe and growth of perturbations (1.3). We summarise the physics of the early Universe in (1.4) and catalogue the energy components that make up the energy budget of the Universe (1.5) before discussing some of the probes cosmologists use to study the Universe (1.6).

Weak Gravitational Lensing forms the basis of the work in this thesis, so we discuss it in some detail (1.7), from the bending of light due to gravity (1.7.1.2) to the weak distortions caused by the projected density field (1.7.2) and some of the statistics used to quantify it (1.7.3 to 1.7.4). Systematic effects, past and future surveys and related probes are treated in (1.7.5 to 1.7.8)

1.1 Cosmology & the ΛCDM paradigm

1.1.1 The Metric

Our observations of the Universe have lead to the perhaps surprising conclusions that it is homogeneous and isotropic, i.e. the Universe has no preferred locations or directions. Clearly on small scales this is very far from the case. Our planet, solar system, galaxy and local group are most definitely the result of highly inhomogeneous structure formation.

On larger scales ($>100h^{-1}$Mpc, where h is the hubble constant and Mpc is a megaparsec, or 3.085×10^{22}metres) the picture is very different. The Cosmic Microwave Background provides the most striking evidence of isotropy- that the Universe looks the same in all directions. Variation across the CMB is of the order $\frac{\Delta T}{T} < 10^{-5}$ (Smoot & et al. 1992), where T is the mean CMB temperature and ΔT is the fluctuation around the mean. Homogeneity - that the Universe looks the

same at every point in space - is harder to prove observationally. Our observation of local isotropy can be coupled with the 'Copernican Principle' that there is nothing special about our location in the Universe. If this is so then all points in space will observe isotropy. It is provable from simple geometry that a space which is, at every point, locally isotropic is globally homogeneous. Note that the reverse is not the case and homogeneity is not a proof of isotropy. For example, a homogeneous universe may be anisotropic due to a global rotation. A complete list of possible relativistic models possible universe exists and is known as the Bianchi classification.

Taken together, the description of our Universe as homogeneous and isotropic is sometimes known as the 'Cosmological Principle.' This may also include the tacit assumption that the laws of physics we observe locally hold on cosmic scales and different epochs in the evolution of the Universe. This simple principle has a number of extremely powerful consequences.

Most importantly we can write down the most general spacetime metric consistent with homogeneity and isotropy, the Friedmann-Lemtaitre-Robertson-Walker (FLRW) metric whose line element is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)[\mathrm{d}r^2 + S^2(r)(\mathrm{d}\theta^2 + \sin^2\theta d\phi^2)] \tag{1.1}$$

where the possible forms of $S(r)$ are

$$S(r) = \{\sin(r), r, \sinh(r)\}, \tag{1.2}$$

r is comoving co-ordinates and t is cosmic time.

We can recast the FLRW metric via a simple change of variables into

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{1.3}$$

where

$$k = \begin{cases} +1 & \text{if } S(r) = \sin(r) \\ 0 & \text{if } S(r) = r \\ -1 & \text{if } S(r) = \sinh(r) \end{cases} \tag{1.4}$$

represents the spatial curvature of the space and governs the geometric properties of the homogeneous and isotropic 3-space corresponding to the hypersurface t =const. $k = +1$ corresponds to positive spatial curvature where parallel light paths will converge; $k = 0$ to zero spatial curvatures equivalent to ordinary flat Euclidean space $k = -1$ to negative spatial curvature and hyperbolic geometry where parallel light paths diverge.

1.1.2 Expansion and Redshift

Arguably the founding event of the modern discipline of cosmology was Edwin Hubble's study (Hubble 1929) of the redshift-distance relation of nearby galaxies using Cepheid variables as standard candles. The discovery of the relation

$$v = Hd \tag{1.5}$$

known as "Hubble's law", which relates the velocity of galaxies to their distance from the observer, has far-reaching consequences. As they are proportional and related by the Hubble parameter, the rate of change of the scale factor, $H = \frac{\dot{a}}{a}$, tells us we live in an expanding Universe. If we wind the clock back we find some time at which the Universe was arbitrarily small, hot and dense. Current best measurements put the time to this "beginning", hence the age of the Universe at 13.69$\pm$0.13billion years (Dunkley *et al.* 2009). Hubble's discovery was the first clue towards the initial conditions of the Universe which we now describe as a Hot Big Bang, an initial singularity from which our Universe was born.

One consequence of an expanding Universe is that light, emitted by some source, will be shifted to a longer wavelength en route to an observer. This *redshifting* can be related to the scale factors at emission and observation

$$\frac{\lambda_0}{\lambda_e} = \frac{a_0}{a_e} \equiv 1 + z \tag{1.6}$$

and defines a parameter, z, called the redshift.

1.1.3 Distances

The proper distance is what we would measure if we could place rulers end-to-end between two objects in the Universe at one particular instant. More useful for cosmography (the measurement of the Universe) is the comoving distance, χ, between two objects which remains constant with the cosmic expansion.

The comoving distance along the line of sight out to a redshift z is defined as

$$\chi(z) = D_H \int_0^z \frac{dz'}{E(z')} \tag{1.7}$$

$$\text{where } E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda} \tag{1.8}$$

where $D_H \equiv c/H_0 = 3000h^{-1}$Mpc is the Hubble distance Hogg (1999). Ω_X are the dimensionless energy densities for different components which make up the total energy budget of the

Universe, namely matter (m), curvature (k) and a cosmological constant (Λ). See section 1.5 below for more discussion of these components and their evolution with redshift. The $E(z)$ formalism is a useful innovation of Peebles (1993). Comoving distance is the fundamental distance measure in cosmography but it is not directly measurable. In practice we define some operational distance measures based on χ.

Angular Diameter Distance

Consider an object of known physical size. We have access to the angle it subtends on the sky but we can define a distance via simple Euclidean geometry as the ratio of an object's physical size to its angular size (in radians):

$$d_A = \frac{l}{\Delta\theta}. \tag{1.9}$$

From the FRW the proper distance along a circle is $l = aS(\chi)$, hence

$$d_A = aS(\chi) = \frac{S(\chi)}{1+z} \tag{1.10}$$

or, in a flat Universe ($S(\chi) = \chi$), simply

$$d_A = \frac{\chi}{1+z}. \tag{1.11}$$

Note we have defined the scale factor today to be unity.

The angular diameter distance can be measured from an observed angular size when we feel confident of the size of an object on the sky, e.g. the acoustic fluctuations in the CMB or baryon acoustic oscillations (BAOs) imprinted on the galaxy distribution on large scales.

Luminosity Distance

In a static, Euclidean Universe a source with absolute luminosity L at a distance d will be observed to have a flux F

$$F = \frac{L}{4\pi d^2}. \tag{1.12}$$

The flux falls off as the surface area of a sphere centred on the source increases. We define the luminosity distance

$$d_L = \left(\frac{L}{4\pi F}\right)^{1/2}. \tag{1.13}$$

To express this in terms of an expanding FRW Universe we should note that the total rate of energy coming through the sphere is modified from the emission rate by two effects. The number of photons coming through the sphere per unit time is reduced by $1/a = 1 + z$ and each photon is also redshifted so has its energy decreased by an additional factor of $1/a = 1 + z$. So the total energy decreases by a factor $(1 + z)^2$ and we can write

$$F = \frac{L}{4\pi(S(\chi))^2}\frac{1}{(1+z)^2}. \tag{1.14}$$

Substituting into eqn. 1.13 we find $d_L = S(\chi)(1+z)$ or, in a flat Universe

$$d_L = \frac{\chi}{a}. \tag{1.15}$$

This distance is applicable to standard candles, most interestingly in a cosmological context to Type Ia Supernovae (see section 1.6.3).

Comoving Volume This is a volume measure in which the number densities of non-evolving objects locked into the Hubble flow are constant with z. It is equal to the proper volume times a factor of $(1+z)^3$. We can write the comoving volume element as

$$\mathrm{d}V_c = D_H \frac{(1+z)^2 D_A^2}{E(z)} \mathrm{d}\Omega \mathrm{d}z. \tag{1.16}$$

1.2 Gravity

Gravity is the most pervasive and important force in our observable Universe. Whether we wish to describe the orbit of a planet, the formation of a black hole or the dynamics of the whole Universe, gravity - the mutual attraction of massive bodies - is the key ingredient. In the next section we briefly discuss the ancestry of our modern theory of gravity. Then in section 1.2.2 we summarise some of the key results of that theory, Einstein's general relativity, before discussing some alternative approaches in section 1.2.3.

1.2.1 Relativity- Galileo, Newton, Einstein

The study of gravitation- why the cannonball falls or the planets orbit as they do, has been a rich seam running through the history of physics and astronomy. Later chapters in this thesis deal with the study of gravity models on cosmological scales.

We perhaps recognise the origin of our modern understanding of gravity in the work of Galileo, the 16$^{\text{th}}$/17$^{\text{th}}$ Century Tuscan astronomer, mathematician and philosopher. It was Galileo who first proposed the principle of relativity, noting that a person in the hold of a smoothly moving ship has no sense of movement, there is no *internal* observation one can make that will distinguish the moving ship from a room at rest. Galileo also proposed something like the equivalence principle when he asserted that the inertial and gravitational masses of an object are identical.

It was a return to some of these ideas that bore great fruit for Einstein in the early 20th century. The intervening period had seen Newton's epoch-making work on gravity- universal gravitation; the inverse square law; equal action and reaction- come to dominate the subject.

There were however some discrepancies between Newtonian theory and observation, foremost of which was the precession of the perihelion of the planet Mercury. It was Einstein's return to the relativity principle, applied to light not in the hold of a ship but in the somewhat less queasy surroundings of a smoothly moving sealed train carriage, which lead to his Special Theory of Relativity, dealing with objects moving at constant velocity close to the speed of light.

It was a reformulation of the equivalence principle, stating that the effects of a gravitational field are equivalent to the effects of acceleration in the absence of a gravitational field that led Einstein to conclude that free-falling objects are undergoing inertial motion. Eventually this produced a more general theory of gravitation, mass and acceleration which we briefly discuss in the next section.

1.2.2 General Relativity

The genius of General Relativity (GR) is its geometrical nature. Doing away with forces acting on bodies, GR instead has that gravitating mass curves spacetime locally. The dynamics of test particles in the curved spacetime accounts for the effects of gravity that we observe. Mass (and energy, they are equivalent) source spacetime curvature through the stress-energy tensor which, for a perfect fluid, takes the form

$$T_{\mu\nu} = (\rho + p)U_\mu U_\nu + p g_{\mu\nu} = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{1.17}$$

where U^μ is the fluid 4-velocity, ρ is the energy density in the rest frame and p is the pressure in that frame. The metric tensor, $g_{\mu\nu}$, describes distances in the spacetime of interest, for our expanding homogeneous and isotropic Universe we use the FLRW metric previously discussed.

The stress-energy tensor is related to the curvature of spacetime via Einstein's Field Equations

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = -\frac{8\pi G}{c^4} T^{\mu\nu} \tag{1.18}$$

where $R^{\mu\nu}$ and R are the Ricci tensor and scalar respectively and describe the curvature.

In Minkowski spacetimes (spacetime analogue of Euclidean space, suitable for Special Relativity) test particles travel in straight lines unless deflected by a force. In the four dimensional spacetime of general relativity we generalise the concept of a straight line to that of a geodesic, which defines the shortest path between two points.

We can use Einstein's field equations to derive the Friedmann and acceleration equations as well as describe perturbations around the flat FRLW background, all summarised in section 1.3.

1.2.3 Modified Gravity Theories

General Relativity is an astoundingly successful theory. Solar system tests from laboratory experiments to laser-ranging using the moon and spacecraft in the outer solar system have confirmed Einstein's theory to order 10^{-13} (Adelberger *et al.* 2003; Schlamminger *et al.* 2008).

On cosmological scales and at distant times the story is not quite as clear cut. We might expect the cosmological principle to hold and allow us to extrapolate local gravitational behaviour to cosmic scales, indeed there has been no detected deviation from GR on any scale. Nevertheless there still exists some room to manoeuvre on these scales as well as strong motivation to explore Modified Gravity (MG) theories which we examine in detail later.

Equally successful as GR in the 20th century was the theory of quantum mechanics, dealing with physics on very small scales. It is clear that, as currently formulated, QM and GR are incompatible. For similar reasons GR breaks down in the extremely hot and dense conditions of the very early Universe and Big Bang, offering further motivation for alternative descriptions of gravity which might unite it with the other three fundamental forces.

The theoretical effort devoted to creating MG theories is impressively varied. Some of the most studied approaches include: *f(R) theories* and other chameleon theories which add a correction term to the Einstein-Hilbert action; *The DGP model* in which gravity acts in a 5D braneworld, the extra dimension weakening gravity as it "leaks-out" of our 4D spacetime and *Degravitation* where gravity acts as a high-pass filter and the vacuum energy "degravitates", possibly explaining the fine-tuning problem. We investigate deviations from GR from chapter 3 onwards. For more background see Durrer & Maartens (2008).

1.3 Dynamics, Perturbations & Structure growth

It is worth reiterating that the FLRW metric relies only on the assumptions of homogeneity and isotropy. It is a purely kinematic description and does not assume a theory of gravity. Einstein's GR relates gravitating mass to local geometry; to derive the dynamics of our Universe we apply Einstein's equations to the FLRW metric. The description of dynamics in an FLRW Universe is essentially an attempt to describe the evolution of the scale factor a.

The combination of the metric and the Einstein equations produces two independent equations

which describe the evolution of the scale factor, the Friedmann equation

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\sum_i \rho_i - \frac{k}{a^2} \tag{1.19}$$

which relates the rate of increase of a to the total energy density of all matter in the Universe, and the acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\Sigma_i(\rho_i + 3p_i). \tag{1.20}$$

The sum over i includes the different matter contributions to the overall energy density of the Universe where i = { matter, radiation, vacuum energy}. The summation assumes the components are non-interacting which is robust for perfect fluids at late times and stems ultimately from the fact that one can sum the energy-momentum tensors for separate, non-interacting components

$$T^{\mu\nu} = \sum_i (T^{\mu\nu})_i \quad .$$

Note that this property does not hold if there exist interactions between components in the dark sector, e.g. dark energy which decays into dark matter.

It is useful to define a critical density from the Friedmann equation

$$\rho_{\rm crit} = \frac{3H^2}{8\pi G} \tag{1.21}$$

in a flat Universe. The critical density allows us to define a dimensionless energy density $\Omega \equiv \frac{\rho}{\rho_{\rm crit}}$. Again, each matter component has its own energy density Ω_i and we can relate the total energy density to the local geometry of the Universe

$$\Omega_{\rm total} = \Sigma_i \Omega_i \begin{cases} > 1 & k = 1 \\ = 0 & k = 0 \\ < 1 & k = -1 \end{cases} . \tag{1.22}$$

A third, useful but not independent, equation can de derived for consistency via energy conservation $\Delta_\mu T^{\mu\nu} = 0$

$$\dot{\rho} + 3H(\rho + p) = 0. \tag{1.23}$$

The assumption that the matter components can be described as perfect fluids means we can relate the energy density and pressure by a single parameter, the equation of state $w = p/\rho$. The separability of the energy-momentum tensor means that energy conservation holds for each component separately and each has their own equation of state w_i. If a component has a constant equation of state the evolution equations are straightforward and the energy density varies with a as

$$\rho(a) \propto \frac{1}{a(t)^{3(1+w)}}. \tag{1.24}$$

1.3.1 Perturbations and Growth of Structure

The expansion of the Universe, the evolution of the scale factor in a FLRW metric under General Relativity and the evolution of matter density with a can be said to describe the *smooth* expansion and evolution of our Universe but one of the most obvious properties of our local universe is that it is very far from smooth. Dark matter haloes, galaxy clusters, even stars and planets are all examples of extremely non-linear inhomogeneity in the local distribution of mass.

The source of the inhomogeneity is gravitational instability. Any initially overdense region will tend to attract more mass via gravity. No matter how small the initial overdensity, matter will continue to accumulate, allowing the growth of structure in the Universe. Schematically this growth stems from the competition between gravity and pressure which slows the accumulation of mass

$$\ddot{\delta} + [\text{Pressure} - \text{Gravity}]\delta = 0 \tag{1.25}$$

where $\delta = \frac{\delta(\mathbf{x})-\bar{\rho}}{\bar{\rho}}$ is the local matter overdensity.

It is often convenient to describe perturbations around the flat FLRW background using the metric in the conformal Newtonian Gauge (Mukhanov *et al.* 1992):

$$\mathrm{d}s^2 = -(1+2\Psi)\mathrm{d}t^2 + a^2(1-2\Phi)\mathrm{d}\mathbf{x}^2 \tag{1.26}$$

Note that this simple version of the line element restricts us to considering scalar perturbations only. This is fine for all work within the scope of this thesis. This gauge can be generalised to vector and tensor modes. A benefit of this gauge is that Ψ corresponds to the Newtonian potential for the acceleration of particles. This gives Ψ an intuitive physical interpretation and, for GR and in the absence of anisotropic stress in $T_{\mu\nu}$, the two potentials are identical.

Our particular theory of gravity specifies the relation between the two metric potentials and between the potentials and perturbations in density, δ and peculiar velocity, v. In GR the metric potentials are equal, $\Phi = \Psi$ and can be related to the gravitational potential via the Poisson equation (Dodelson 2003)

$$k^2\Phi = 4\pi G\rho_m a^2\delta. \tag{1.27}$$

There are three distinct regimes which describe the growth of perturbations. Outside the horizon (corresponding to comoving scales greater than the Hubble length $(1/aH)$) gravitational potential grows at a fixed rate $\propto a^2$ during radiation domination and $\propto a$ during matter domination. We say a scale "enters the horizon" when its scale length is comparable to the horizon scale,

$k = (aH)_{enter}$. During a decelerating phase in the evolution of the Universe the Hubble scale grows in comoving coordinates, so new scales are constantly entering the horizon. Their subsequent behaviour depends on whether a mode enters before or after the epoch of matter-radiation equality. If a mode has a scale smaller than that of the Hubble scale at matter-radiation equality ($k < k_{eq}$) then it enters in the radiation dominated era. The growth of these scales is suppressed during radiation domination. After matter-radiation equality all modes inside the horizon grow as $\propto a$.

Conversely, modes with $k > k_{eq}$ will enter the horizon after matter-radiation equality. They have been growing as $\propto a^2$ during radiation domination and $\propto a$ during matter domination which continues inside the horizon (Dodelson 2003; Carroll 2001). The relations between growth and scale factor are found by solving the growth equation, see eqn. 1.29 below. There exist two solutions: growing and decaying modes, we are interested in the growing modes.

Taken together these effects create the the characteristic shape of the matter power spectrum, shown in Fig. 1.2. Small-scale modes have experienced suppression inside the horizon during radiation domination so have lower amplitudes than large scale modes which entered after matter-radiation equality. The scale of the turnover is that of the mode which enters the horizon at matter-radiation equality. After this time the shape of the matter power spectrum is frozen-in and scale-independent growth proceeds (ignoring non-linear effects at small scales) because radiation is now too weak to impede gravitational clustering.

These processes are sometimes written as

$$\Phi(\mathbf{k}, z) = \Phi_p(\mathbf{k}) \times T(k) \times D(z) \tag{1.28}$$

where $\Phi_p(\mathbf{k})$ is the primordial potential imprinted at the end of inflation, $T(k)$ is the transfer function describing the evolution of perturbations through horizon-crossing and matter-radiation equality, and $D(z)$ is the growth function which describes the scale-free linear growth via

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\rho\delta. \tag{1.29}$$

This result comes from linear perturbation theory, see e.g. Liddle & Lyth (2000).

1.4 The Early Universe

1.4.1 The Big Bang

We have mentioned that the expansion of the Universe and the dynamical equations make it plain that the early Universe began in an extremely small, hot and dense state. Called the “primaeval

atom" by Lemaître, we now know this state as the Big Bang, an initial singularity from which the whole Universe sprang. It is a common misconception to imagine the Big Bang as an explosion from which the Universe expanded out. This is misleading as there was no "place" where we could point to a Big Bang explosion occuring and the Universe is certainly not expanding out into anything. Rather the Big Bang occurred everywhere equally and the expansion comes from the fact that the "distance" between objects, i.e. the scale factor, is increasing.

The cause of the Big Bang, if we can talk of such a thing, is the subject of much feverish speculation but even the very first moments after the Big Bang are inaccessible to our current physics toolbox, requiring a full theory of quantum gravity to come to grips with them. We would perhaps expect the Universe to expand and cool smoothly after the Big Bang. Indeed it behaves like this for most of its history but there is an unusual intervening epoch described in the next section.

1.4.2 Inflation

We have introduced many of the ingredients of the standard picture of the Universe and its evolution. By the 1980s observation of the CMB had firmly established the Hot Big Bang paradigm and largely ended the internicene warfare with proponents of the Steady State theory which had dogged the 1960s and 70s. However, long before the discovery of accelerating expansion, there remained a number of niggling problems which seemed inconsistent with a Universe populated by radiation and gravitating mass evolving via General Relativity in the aftermath of an initial hot, dense singularity. These included:

The Flatness Problem We can rewrite the Friedmann equation as

$$|\Omega(a) - 1| = \frac{|k|}{a^2 H^2} \tag{1.30}$$

with derivative

$$\frac{d|\Omega|}{da} = (1 + 3w)\frac{\Omega(\Omega - 1)}{a}, \tag{1.31}$$

where Ω is the combine dimensionless energy density of all the components in the Universe and w is the combined equation of state Carroll (2001). If, as observed, the total density today is very close to unity, $\Omega(a_0) = 1$, then we are sitting very close to a repeller solution to this equation for a Universe dominated by ordinary matter and radiation ($w > -1/3$). This means that any slight deviation from flatness will evolve as $|\Omega(a) - 1| \propto t^{\frac{2}{3}}$ towards an increasingly curved geometry.

To match the observed value of Ω_{k0} then Ω in the very early Universe must lie in the (conservative) range

$$0 \leq 1 - \Omega \leq 10^{-60}. \tag{1.32}$$

This presents a difficult fine-tuning problem as even small changes in the early value will lead to a wildly different cosmology today.

The Horizon Problem We have previously noted that the Universe, as evidenced by the CMB, is isotropic to a high degree on large scales. The Universe at recombination was hot and in thermal equilibrium so this may not seem a particularly surprising result. The CMB temperature contrast has been measured as $\frac{\Delta T}{T} \sim 2 \times 10^{-5}$ at separations of 10° on the sky. The problem arises because the horizon distance on the sky is of order 1°. Somehow regions that have never been in causal contact with each other and were widely separated at the surface of last scattering are in a state of thermal equilibrium. This apparent communication between causally disconnected regions calls into doubt our description of the evolution of the Universe.

The Relics Problem As the temperature of the Universe cools following the Big Bang, an originally united fundamental force is thought to "freeze out" through a series of phase transitions into the four known fundamental forces we see today. Description of this process is the goal of grand unified theories (GUTs). The symmetry breaking which occurs during the phase transitions is thought to produce relic topological defects, particularly magnetic monopoles, in large numbers.

To date, no magnetic monopole has been observed. Assuming they exist, this means that monopoles must be particularly rare- very much in contradication to the best predicitons of GUTs.

The Solution One suggested theory, which has now become a standard part of our cosmological picture, was developed through the 1980s (Guth 1981) and elegantly solves each of the problems listed above. The theory, really a competing family of theories, is called Inflation and posits that the very early Universe underwent a period of accelerating expansion ($\ddot{a} > 0$). Just as with the dark energy dominated epoch, this requires an equation of state $w < -1/3$. The simplest inflation models populate the accelerating Universe by a single scalar field, the inflaton, which powers the expansion. When inflation ends, the inflaton decays into the particles and radiation we see today, thermalising in a process known as reheating.

The period of accelerating expansion increases the denominator in eqn. 1.30, pushing Ω_K so

close to zero that any deviations from flatness in the post-inflationary epoch are negligible.

In addition, the inflationary epoch quickly expands small regions of space to huge sizes, well outside our current horizon. Sources we see as widely separated on the CMB sky were very close before inflation, allowing causal contact and the observed thermal equilibrium.

Similarly the rapid enlargement of small spatial volumes massively dilutes the density of relic topological defects, including magnetic, monopoles. Enough inflation makes it extremely likely that there would be none of these exotic relics in the volume of space which happens to be accessible to our observations.

The solution of each of the problems that provided the need for inflation requires a certain amount of accelerated expansion, generally stated as N_e, the number of "e-folds" for which inflationary evolution lasted. An e-fold is the amount of time taken for the scale factor to increase by a factor of e. Current theories require $N_e > 60 - 70$, which requires a time of roughly 10^{-34}s.

An additional benefit of inflation, not part of the original motivation, is now perhaps its most important property. Inflation vastly enlarges quantum mechanical fluctuations to cosmic scales, leaving behind a spectrum of remnant density perturbations. We have already noted that the CMB's extreme temperature uniformity suggests that the early Universe was homogeneous. Yet we see very inhomogeneous Large Scale Structure in today's Universe. The perturbations left after inflation are an excellent candidate to provide the "seeds" of the structure formation which has occured. The fluctuations are predicted to be approximately scale-free, in good agreement with our observations of LSS.

1.4.3 The Post-Inflation Universe

After the end of inflation the Universe continues to expand and cool, slowed down by the gravitation of matter (at least until the dark energy dominated epoch). The early Universe was dominated by radiation. Temperatures were so high that photons could easily ionise the neutral hydrogen present. The Universe was filled with plasma and the free electrons made it opaque to light.

A point beyond the scope of this thesis but interesting to note is what is known as the problem of *baryogenesis*. In the very early Universe light elements were generated (during reheating according to inflation). Our standard model of particle physics suggests matter and anti-matter should be generated in equal amounts. This cannot be the case or they would have annihilated each other absolutely and none of the matter we see today would have remained to collapse and form structure. There is an unknown asymmetry in the production of baryons in the early Universe which we hope shall be explained by some future model of high energy physics. Andrei Sakharov

gave thought to this problem and proposed three conditions that must hold in order to convert a matter/anitmatter symmetric universe into a matter dominated one. These Sakharov conditions are: baryon number violation; CP violation, and a departure from thermal equilibrium.

As the Universe cooled, photons lost energy until they were no longer able to ionise hydrogen. Protons and free electrons combined to form atoms. This is known as the epoch of *recombination*. Shortly afterwards photons decoupled from the neutral plasma and were able to free-stream. For more details see section 1.6.2.

1.5 Energy Components of the Universe

We have described how the curvature of speacetime in GR is sourced by the stress-energy tensor, $T_{\mu\nu}$. To solve the Einstein equations for the actual expansion history of the Universe we need some knowledge of the components which make up $T_{\mu\nu}$.

Today independent constraints from multiple cosmological probes have converged on a concordance cosmology of $\sim$5% baryons, $\sim$20% dark matter and $\sim$75% dark energy. We are relatively certain of the values of these relative densities but we are much less sure why they should take the values they do and we certainly don't have a coherent theory to explain the origin and size of each component. More detail of the properties of each matter component are detailed below.

1.5.1 Baryonic Matter

Baryonic matter is what we sometimes rather anthropomorphically refer to as "ordinary" matter, made from atoms and their constituents which we are most familiar with. This matter makes up the stars, planets, gas and dust in the Universe, whether immediately visible or not and of course we ourselves are made of baryonic matter.

Strictly baryons are particles made up of three bound quarks (e.g. neutrons, protons). Cosmologists use baryon more loosely, embracing electrons and other leptons. In fact the vast majority of the mass from this type of matter *is* in the particle physicist's baryons.

The primordial baryons distribution was formed in the first 100s after the Big Bang in a process known as Big Bang Nucleosynthesis (BBN). The physics involved is well understood, being essentially the same as that powering thermonuclear processes in stars. The abundances of elements (all very light) produced in BBN is 74% Hydrogen (^{1}H), 26% Helium (^{4}He), 3×10^{-5} % Deuterium (^{2}H), 10×10^{-9} % Lithium (^{7}Li).

We can relate the energy, E, momentum, d, and rest mass, m, of baryons

$$E^2 = d^2 + m^2. \tag{1.33}$$

As the Universe evolves and cools, the energy and momenta of baryons become smaller with respect to their rest mass, i.e. $d \ll m$. We say they are non-relativistic and we can also make the approximation that they form a pressureless fluid, $p \equiv 0$.

Solving the fluid equation for a pressureless fluid gives $\rho \propto 1/a^3$, meaning that their density drops in proportion to volume.

Latest studies from the CMB (Dunkley *et al.* 2009) put $\Omega_b = 0.0441 \pm 0.0030$, meaning that baryons make up under 5% of the total energy of the Universe.

1.5.2 Dark Matter

In fact, most of the mass in the Universe is not made up of "ordinary" baryons, rather it is non-baryonic matter called Dark Matter becuase it interacts extremely weakly, if at all, via the electromagnetic force. It doesn't reflect or absorb light and is therefore invisible to our standard detectors, hence "dark."

The evidence for Dark Matter dates back to the 1930s when Zwicky noticed anomolous results in galaxy velocity measurements from the Coma cluster which could not be explained by the visible mass (Zwicky 1933). Later measurements of the stellar rotation curves in spiral galaxies found a similar result (Rubin *et al.* 1999). It was expected, based on the visible mass, that velocity would vary with radius from the galactic centre as $v_{\rm orb} \propto r^{-1/2}$ but measurements showed velocity remaining constant with radius. This could be explained by the existence of a larger halo of invisible mass within which the galaxy was embedded.

Attempts were made, notably by the Modified Newtonian Dynamics (MOND) theory (Milgrom 1994) to explain these results by modifications to General Relativity rather than the existence of dark matter. There are parallels with the explorations of MG discussed in chapter 3 onwards in this thesis but the basic MOND idea was dealt a death blow by the discovery of the Bullet Cluster 1E 0657-56 (Clowe *et al.* 2006), shown in Fig. 1.1. This is actually two clusters which collided 150million years ago. Clusters certainly contain visible mass in the form of galaxies and, predominantly, Intra-Cluster Gas (ICG). ICG has a much larger collision cross-section than galaxies so, upon collision, galaxies were able to pass by each other relatively unperturbed. In contrast the ICG felt the impact strongly and lost much of its momentum. Images of the cluster show a central region of shocked ICG with the unperturbed galaxies travelling onwards on either side.

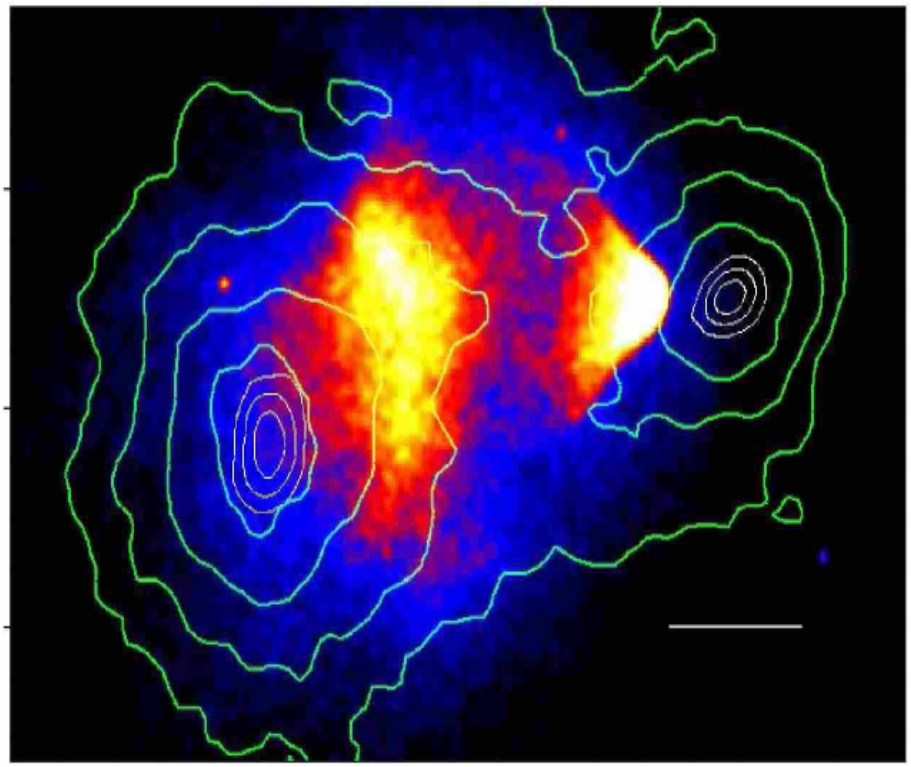

Figure 1.1: Chandra 500ks X-ray image of the "Bullet Cluster" 1E 0657-56. Green contours show the weak lensing κ mass reconstruction with the outer contour level at $\kappa = 0.16$ and increasing in steps of 0.07. There is definite offset between the X-ray emitting shocked ICG in the centre and the main mass concentrations which have passed through each other. Clowe *et al.* (2006)

When gravitational lensing was used to map the mass distribution on the cluster it observed large mass concentrations centred on the galaxy populations but much larger than the visible mass could account for. As the ICG has been stripped from the galaxy populations there is no other baryonic matter present to account for this additional mass. It must be dark matter.

We have referred to DM as non-baryonic but there are potential and actual candidates for baryonic dark matter, referred to as MACHOs (MAssive Compact Halo Objects) these include non-luminous massive objects such as primordial black holes or brown dwarfs. While these do exist, surveys (Alcock *et al.* 2000) show they are far too few to account for the observed dark matter contribution to the energy density of the Universe.

Dark Matter itself can come in different varieties- the most important distinction being between Hot and Warm Dark Matter (species which remain relativistic until late or relatively late times) and Cold Dark Matter which become non-relativistic in the early Universe. As noted below in 1.5.4 neutrinos are a form of HDM and the only observed DM particle. However, the vast majority

of DM must be CDM because HDM & WDM free-stream out of matter overdensities, removing mass from clusters and suppressing growth on small scales in ways contradicted by evidence from LSS.

What we are left with is $\Omega_{\mathrm{cdmh}^2} = 0.1099 \pm 0.0062$ (Dunkley *et al.* 2009), which corresponds to $\sim 21\%$ of the energy density of the Universe is made up of a type of exotic non-baryonic, weakly interacting particle which has never been observed and fits none of the properties of any particle in the standard model of particle physics. Current experiments at the Large Hadron Collider offer some hope for the discovery of a suitable candidate of which there are several proposed including the lightest stable supersymmetric particle (Feng 2010).

For cosmological purposes we can sum the contributions from baryons and CDM to produce the total matter energy density

$$\rho_m(z) = \rho_b(z) + \rho_{cdm}(z). \tag{1.34}$$

The equation of state for "dust", the term for pressureless matter composed of non-relativistic particles, is $w_m = 0$ so the matter density evolves as $\rho_m(z) \propto 1/a(z)^3$.

1.5.3 Dark Energy

With evidence that the geometry of the Universe is very close to flat but only 25% of the necessary energy budget accounted for by baryons and CDM there remained increasing speculation about the nature of the remaining contribution. In addition the age of the Universe computed in a purely CDM universe was actually less than the known age of some of the oldest stars, found in the magellanic clouds Jaffe (1996). These results meant a "concordance cosmology" with a non-zero cosmological constant was beginning to emerge.

It was in this environment that two teams (Riess *et al.* (1998), Perlmutter *et al.* (1999)) studying supernovae as standard candles (see section 1.6.3) produced results showing that the expansion of the Universe is not slowing down, as we would expect for GR acting on standard matter, but actually accelerating.

An accelerating expansion requires $\ddot{a} > 0$ and, from the acceleration equation, $(\rho + 3p) < 0$ i.e. an equation of state $w < -1/3$ from a *negative pressure* fluid.

In the presence of a cosmological constant we can generalise the Friedmann equations as:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\sum_i \rho_i - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{1.35}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i (\rho_i + 3p_i) + \frac{\Lambda}{3}. \tag{1.36}$$

These are then more correctly called the Friedmann-Lemaitre equations.

This discovery confirmed a renewed interest in the Cosmological Constant, Λ, originally introduced by Einstein into his field equations to produce a static solution. The cosmological constant can be subsumed into the stress-energy tensor and becomes another component of the energy density of the Universe with $w_\Lambda = -1$. This contributes an energy density

$$\rho_\Lambda = \rho_{\Lambda,0} = \frac{\Lambda c^2}{8\pi G} = \text{const.} \quad . \tag{1.37}$$

The energy density being uniform and smooth throughout space neatly identifies the cosmological constant with a vacuum energy from "empty" spacetime.

Upon Hubble's discovery of the expansion of the Universe, Einstein unceremoniously dropped the cosmological constant but it has properties ideally suited to power the accelerating expansion. Nevertheless its properties remain obscure, earning it the name Dark Energy to signify our ignorance. One major problem is the fact that we have no fundamental theory that motivates the size of the cosmological constant. Best estimates from quantum mechanics (Carroll *et al.* 1992) give a vacuum energy on the scale $\rho_{vac}^{\text{QM}} \sim M_p^4 \sim (10^{18}\text{GeV})^4$, where M_p is the Planck mass. Embarrassingly this compares with the observed value as

$$\rho_{vac}^{\text{obs}} \equiv 10^{-120}\rho_{vac}^{\text{QM}}, \tag{1.38}$$

over 100 orders of magnitude out. This is known as the *fine tuning problem*. On top of this we know that the ratio of the matter and dark energy densities should scale as $\frac{\Omega_\Lambda}{\Omega_m} \propto a^3$, meaning that dark energy was negligible at early times and comes to dominate in the late Universe. There exists only a brief epoch where their values are comparable but that seems to be exactly where we find ourselves. This is known as the *coincidence problem*. Any well motivated theory of dark energy should explain both these open questions if we are to avoid anthropic arguments.

$w = -1$ is true for a pure cosmological constant but much effort in modern cosmology goes into constraining the behaviour of dynamical dark energy models which allow the equation of state to vary with redshift. There are several motivations including solving the coincidence problem and the hope that understanding the dynamics of dark energy might provide clues towards a theory of its origins. As DE only becomes important in the late Universe the most common parameterisation is a first order Taylor expansion

$$w(a) = w_0 + (1-a)w_a \tag{1.39}$$

where a cosmological constant is recovered for $w_0 = -1$, $w_a = 0$.

There exist alternative models for explaining the observed accelerating expansion of the Universe without recourse to dark energy. If we happen to live in the centre of a cosmic void we would expect to observe similar results even in a Universe where the expansion is slowing. This is related to backreaction theories charting the effects of inhomogeneities on the homogeneous expansion of the Universe. It comes with the heavy price of abandoning the Copernican principle.

Another approach is to modify General Relativity in such a way that the accelerating expansion can be achieved in a Universe only populated by CDM. This approach is discussed in detail from chapter 3 onwards.

1.5.4 Neutrinos

Neutrinos have the distinction of being the only type of dark matter positively detected in laboratory experiments. Originally postulated by Pauli in 1930 to conserve energy and momentum in β-decay, neutrinos were assumed to be massless in the standard model of particle physics. The discovery of oscillations between neutrino flavours in the detected solar neutrinos was the first evidence of physics beyond this standard model as it required neutrinos to have (at least) three finite mass eigenstates.

As well as being produced in supernovae core-collapse, a large flux of neutrinos was created during the Big Bang, leading to the prediction of a "cosmic neutrino background." Neutrinos decouple from other particles early because they only interact via the weak force. After the Universe cools to $\sim$1MeV neutrinos propagate freely through the Universe. Neutrinos are very light and remain relativistic until relatively late times. Relativistic neutrinos are not captured by the gravity wells of large-scale structure, they "free stream" out of over-densities, reducing the mass of large-structures (Bond *et al.* 1980). The exact sum of the masses of the neutrino eigenstates affects their free-streaming length and the time at which we become non-relativistic so the study of LSS and the CMB can place bounds on the neutrino mass. Neutrino mass density is related to the sum of the neutrino mass eigenstates via:

$$\Omega_\nu = \frac{\Sigma_i m_{\nu_i}}{93.14 h^2 eV} \ . \tag{1.40}$$

This equation relates the energy of massive neutrinos as a species to their temperature and assumes that the transition of neutrinos from relativistic to non-relativistic behaviour occurs at $T_\nu \sim m_\nu$. See Dodelson (2003) for a full derivation.

Neutrinos interact only extremely weakly via the electromagnetic force, making them a form of dark matter. As they are relativistic until late times they are termed Hot Dark Matter (HDM)

to differentiate them from Cold Dark Matter (CDM) which decouples at early times. Studies of LSS show that the bulk of the dark matter in the Universe must be in the form of CDM, latest constraints give $\Omega_\nu < 7 \times 10^{-3}$, equivalent to a combined neutrino mass of $\Sigma_i m_{\nu_i} < 0.28$eV (Thomas *et al.* 2010).

1.5.5 Radiation & Curvature

Radiation has an equation of state $w_r = 1/3$ corresponding to $\rho_m(z) \propto 1/a^4(z)$. Qualititavely we can think of this as the dilution with volume experienced by matter plus an extra (1+z) due to the redshifting of photons in an expanding universe. Apart from in the very early Universe, the strong dilution with a means that the radiation contribution to the overall energy density is very small, with $\Omega_{r,0} \ll 1$. It is usually a very sensible approximation to neglect radiation.

In general a positive or negative curvature will also contribute to the stress-energy tensor and the total energy density with equation of state $w_c = -1/3$. Recent observations point to a Universe incredibly close to flat geometry $\Omega_k = 0$ so we will neglect the contribution of curvature to the energy contents of the Universe and write all geometric terms as they would be in a perfectly flat universe.

In general we can approximate the overall energy density as the combination of matter, and dark energy, $\Omega = \Omega_m + \Omega_\Lambda = \Omega_{cdm} + \Omega_b + \Omega_\Lambda$.

1.6 Cosmic Probes

Modern cosmology has developed a number of complementary probes of both the expansion history of the Universe and the growth of structure from gravitational collapse of initial perturbations. In this section we discuss some of the statistics that describe the formation of large scale structure at late times, the Cosmic Microwave Background which offers us a snap-shot of the very early Universe and Supernovae which provide tests of the geometry of the Universe to high redshift. Weak Gravitational Lensing is a relatively new probe with great potential. The use of WGL forms the bulk of this thesis so it is treated in more detail in the next section.

1.6.1 LSS

We have already discussed how the inhomogeneous structure we see in the Universe was seeding from initial fluctuations (produced during inflation) which grow under gravitational instability. The growth of large scale structure (LSS) is a powerful probe of cosmology.

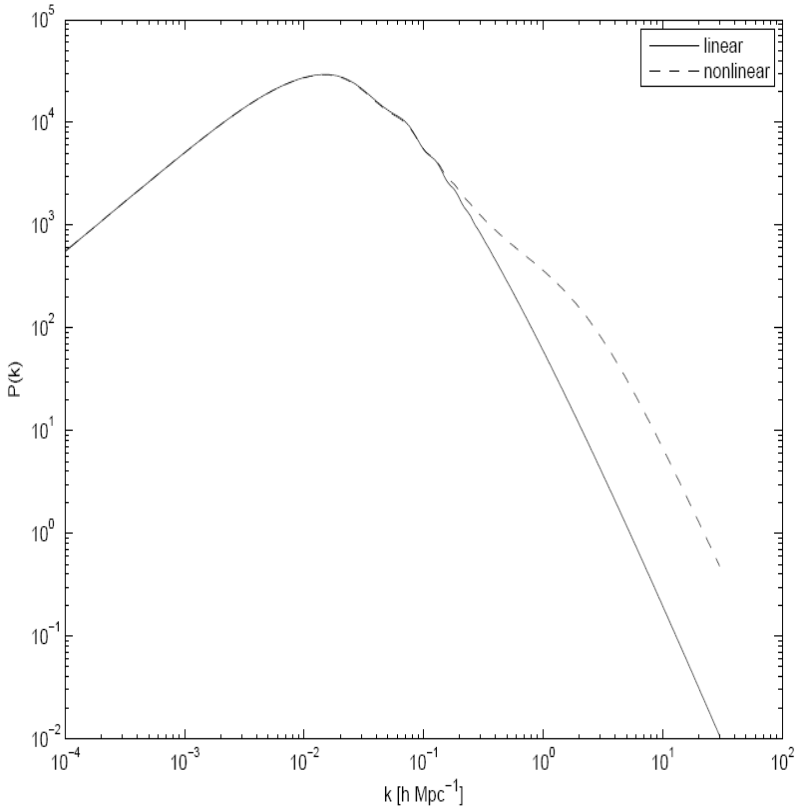

Figure 1.2: ΛCDM power spectrum, normalised to the local abundance of galaxies, for a flat universe, $\Omega_m = 0.25$, $\Omega_b = 0.05$, $\sigma_8 = 0.8$, h = 0.7. Solid line shows the linear power spectrum, dotted line shows the non-linear power spectrum according to the fitting function of Smith *et al.* (2003).

To access this cosmological information it is necessary to treat the complexities of LSS statistically. The most common approach is the two-point correlation function of the matter overdensity, δ, and its Fourier Transform, the Matter Power Spectrum:

$$< \delta(\mathbf{k})\delta^*(\mathbf{k}') >= (2\pi)^3 P(k)\delta_D(\mathbf{k} - \mathbf{k}') \tag{1.41}$$

where δ_D is the Dirac delta function and $< ... >$ is the ensemble average of many realisations of the field at a given separation, **k**. Power spectra uniquely define Gaussian random fields, non Gaussian information is lost to them but can be regained through higher order statistics, see Takada & Jain (2004) for some discussion of 2pt and 3pt statistics in a cosmic shear context.

The power spectrum is a general statistic useful for describing the matter perturbation field but we need particular observable tracers to make use of it in practice. We describe the cosmic shear power spectrum in section 1.7.3. Another common approach is to treat visible galaxies as tracers

of the underlying (predominantly dark-) matter distribution. The baryons that form galaxies should sit roughly at the centre of the gravitational potentials of the total mass so 3D measurements of galaxy position (angular position on the sky plus redshift) from surveys such as the Sloan Digital Sky Survey (SDSS) (Abazajian *et al.* 2009) can be used to construct the 3D matter power spectrum.

In practice galaxies are biased tracers of the underlying density field and the galaxy power spectrum, $P_g(k,z)$ related to the matter power spectrum by

$$P_g(k,z) = b_g^2 P_\delta(k,z) \tag{1.42}$$

where b_g is the galaxy bias. Ignorance of the exact form of the galaxy bias is a major hindrance to our use of galaxy surveys for cosmology, especially on small scales (Kaiser 1988).

Often b_g is assumed to be a constant (linear bias) or vary linearly with redshift but, in general, it is a function of scale and redshift, $b_g(k,z)$. There are a number of different motivated bias models, including the Halo model (Cooray *et al.* 2000) which assumes that mass is concentrated in dark matter haloes with a particular concentration of galaxies defined by the Halo Occupation Distribution. The power spectrum can then be broken into two terms, the 1-halo term due to correlations between galaxies in the same halo and the 2-halo term due to correlations between galaxies occupying separate halos.

A particular power spectrum must be normalised. In this thesis we follow the σ_8 normalisation where the power spectrum is normalised with respect to the variance σ_R^2 of the mass contrast smoothed over a sphere of radius $R = 8h^{-1}$Mpc:

$$\sigma_8^2 = \frac{4\pi}{(2\pi)^3}\int \tilde{W}^2(k)P(k)k^2 dk \tag{1.43}$$

$$\tilde{W}(k) = \frac{3}{(kR)^3}[\sin(kR) - kR\cos(kR)] \tag{1.44}$$

where $\tilde{W}(k)$ is the window function, a Fourier Transform of a 3D sphere.

On small scales matter clustering becomes highly nonlinear, $\delta > 1$ and our standard power spectrum analysis is no longer sufficient to describe the effects at small scales. A number of fitting formulae based on N-body simulations have been produced as non-linear corrections to the matter power spectrum. In this thesis we use the fitting formula of Smith *et al.* (2003) which is based on the halo model approach calibrated from n-body simulations.

1.6.1.1 BAOs & RSDs

Beyond the amplitude or slope of the matter power spectrum, its shape information contains imprints of additional cosmological information.

The "wiggles" at scales smaller than the turnover in the matter power spectrum are called Baryon Acoustic Oscillations (BAOs) and emerge from the physics of the photon-baryon fluid in the radiation-dominated era. As the collapse of structure is halted by radiation pressure, oscillations are generated in the photon-baryon fluid. After baryons decouple, they preferentially attract dark matter on scales of the oscillation, imprinting their particular over- and under-densities on top of the overall P(k). Determining the shape of the wiggles can effectively measure the angular diameter distance and act as a probe of cosmology.

The power spectrum is a 2pt statistic that measures the separation of points tangential to the line of sight but also separation along the line of sight, i.e. redshift. We have mentioned that galaxies are often used as a tracer of the underlying mass distribution. Galaxies are embedded in the Hubble flow of the Universe but they also have their own peculiar velocity on top of the general expansion which responds to the local gravitational potential. These peculiar velocities affect the redshifts we measure for particular galaxies. A peculiar velocity towards the observer acts like a blueshift and one away from the observer acts like an additional redshift. These effects are known as *redshift space distortions*. On small scales, e.g. random motions within a galaxy cluster, elongates structures along the line of sight in what is called the finger-of-god effect. On large scales galaxies are collapsing into overdensities. This produces the opposite effect, squashing structures along the line of sight.

We look at the application of these tools to spectroscopic surveys in combination with WGL in chapter 5.

1.6.2 CMB

The very early Universe was opaque. After the Big Bang (and the inflationary epoch) conditions were extremely hot and dense, atoms were ionised. The free electrons coupled with the photons through Compton scattering. As the Universe expanded it cooled sufficiently for neutral Hydrogen to form, this is the epoch of recombination which made the Universe transparent for photons. The point at which photons were able to free stream is called the surface of last scattering and it is from here that the Cosmic Microwave Background we see today was emitted.

As such the CMB represents a snapshot of the Universe at the epoch of recombination, $z \sim 1100$. The CMB radiation has cooled since emission and has been measured to be a near perfect blackbody spectrum with temperature T = 2.725 +/- 0.001 K (Mather *et al.* 1999).

We have previously noted how isotropic and uniform the CMB is with fluctuations on the order of $\frac{\Delta T}{T} \sim 10^{-5}$. The small fluctuations or temperature anisotropies in the CMB are a very powerful

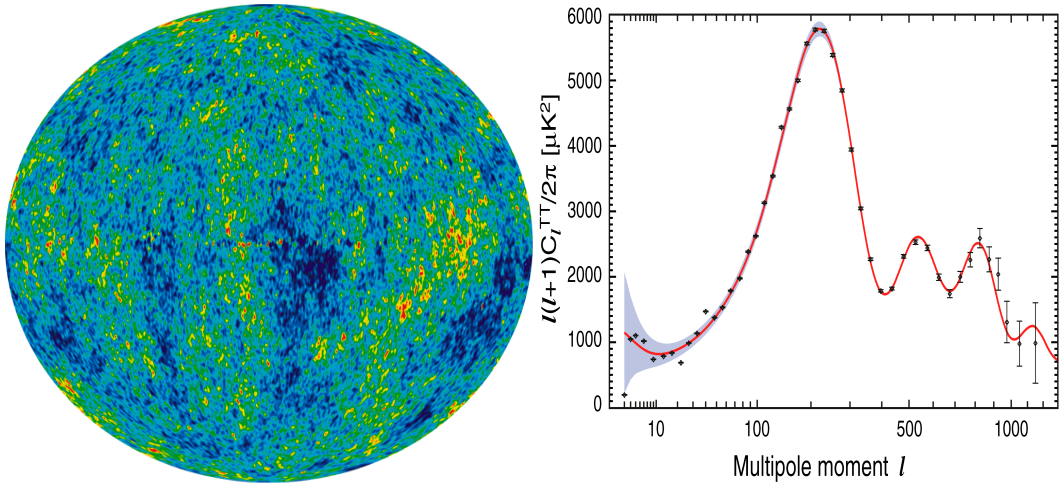

Figure 1.3: *Left panel*:The 7-year temperature (TT) power spectrum from WMAP. The plotted errors include instrument noise. The gray band represents cosmic variance. The points are binned in progressively larger multipole bins with increasing ℓ. Larson *et al.* (2011). *Right panel*: The Internal Linear Combination Map. A weighted linear combination of the five WMAP frequency maps. The weights are computed using criteria which minimize the Galactic foreground contribution to the sky signal. The resultant map provides a low-contamination image of the CMB anisotropy. WMAP Science Team.

cosmological tool. We usually describe them by decomposing their angular correlation function into spherical harmonics:

$$\frac{\Delta T}{T}(\mathbf{r}) = \Sigma_{l=0}^{\inf}\Sigma_{m=-l}^{l} a_{lm}^{T} Y_{lm}(\theta,\phi). \tag{1.45}$$

Anisotropies in the CMB can be broken into two types- Primary Anisotropies imprinted at or just before the surface of last scattering and Secondary Anisotropies generated afterwards.

Primary Anisotropies include *acoustic oscillations* of the photon-baryon fluid due to the competition between gravity and radiation pressure which are imprinted in the CMB at last scattering. The first acoustic peak in particular is strong and related to the sound horizon scale. This can act as a standard ruler and allow us to calculate an angular diameter distance sensitive to cosmology.

CMB photons are also *polarised* due to scattering off free electrons during recombination and the polarisation power spectrum and temperature-polarisation power spectrum offer additional statistical tool to track physics of the CMB and remove systematic effects.

If we imagine the moment of recombination (in reality it takes some short time to become complete), some photons become able to free stream while trapped in local matter overdensities. It will have to climb out of the potential well to escape and free stream. Overdense regions will

appear as cold spots on the CMB sky. The opposite is true for local underdensities. This adds power to the CMB on large scales and is known as the *Sachs-Wolfe effect*.

As mentioned, recombination was not instantaneous. Finite time recombination allows *Silk damping*, washing out anisotropies on small scales and producing a characteristic exponential damping tail.

Secondary anisotropies are sensitive to post-recombination physics, as such they are useful for breaking parameter degeneracies and learning about late Universe phenomena such as dark energy. They include the *Integrated Sachs-Wolfe effect* due to the fact that propagating photons blueshift as they fall into potential wells and redshift as they leave them. In a Universe with a time-independent potential the net effect is zero but in a dark energy dominated Universe the potential varies and the photon's wavelength is altered. The effect is integrated because it depends on all the potentials along the line of sight.

The *Sunyaev-Zel'dovich Effect* is caused by the inverse Compton scattering of photons by ionised gas. It is frequency dependent and can be detected by measurements at different wavelengths. It is a useful tool for cluster detection.

CMB photons will be *gravitationally lensed* by large scale structure as they propagate to the observer. This effect is discussed in section 1.7.8.

1.6.3 Supernovae

Type Ia supernovae are some of the most luminous objects in the Universe. White Dwarfs are the standard remnants of most stars which do not go supernova at the end of their life-cycle. As the remnant core collapses under gravity the density increases until electron degeneracy pressure is sufficient to halt the collapse. The physics of white dwarf collapse and the Pauli exclusion principle places a strong upper bound on the mass of white dwarfs, the Chandrasekhar Limit $1.4M_{\odot}$.

In binary systems, white dwarfs can accrete mass from a companion star. As the white dwarf accretes mass, its core heats up. Eventually, as it reaches the Chandresekhar limit, it will reach the temperature sufficient for carbon fusion and a thermonuclear explosion will occur. This violent event is the origin of the Type Ia supernova- so powerful that it often outshines the whole of its host galaxy.

The uniform composition of white dwarfs and the relatively well-understood physics of the explosion makes type Ia SNe cosmological standard candles. Comparing their apparent magnitude to theoretical uniform intrinsic luminosity (abs. Mag. -19.5) allows us to measure the luminosity distance to the SNe. Combined with a spectroscopic analysis of the SNe light we can test the

distance-redshift relation.

Strictly speaking type Ia SNe are "standardisable" rather than standard candles. Small differences in the composition and local properties of Sne will produce different intrinsic luminosities but there is a well defined correlation between the peak luminosity and the shape of a SNe light curve which allows local differences to be corrected for.

It was through the study of distant type Ia SNe that allowed several groups (Riess *et al.* (2001), Perlmutter *et al.* (1999), Schmidt *et al.* (1998)) to constrain the expansion history of the Universe and provide the first evidence for the acclerating expansion of the Universe at late times.

1.7 Weak Gravitational Lensing

This section introduces weak gravitational lensing as a cosmological probe. We start with the bending of light by mass, present the cosmic shear formalism and derive the shear power spectrum before discussing practical shear measures, systematic effects and related probes.

1.7.1 Gravitational Lensing

1.7.1.1 Newton to Einstein

Newton himself thought about the bending of light by massive bodies in around 1704 as part of his own epoch-making endeavours in both optics and gravitation. By the end of the 18th Century a new generation including Cavendish ($\sim$1784) and von Soldner (1804) had computed the bend angle for a light ray under the influence of a point mass within a wholly Newtonian framework.

A century later, as Einstein was developing his Special theory of Relativity (Einstein 1916), he re-derived the Newtonian bend angle via the equivalence principle. A subsequent in depth analysis produced the first deviation from Newtonian results- an extra factor of two in the bend angle due to the curvature of the spacetime metric.

It was the spectacular confirmation of this prediction by the Eddington eclipse expedition to Principe (1919) (Dyson *et al.* 1920) that propelled Einstein, and to a lesser extent General Relativity, into the forefront of the public imagination. In the mid-1930s Einstein ruled out the possibility of observing an obviously lensed object whose shape or position had clearly been distorted by an extra-solar lens.

It was Zwicky who contradicted the pessimistic assertion, predicting the very next year (1937) that galaxies should act as lenses for background sources, producing multiple images and magnifying distant galaxies which would otherwise be invisible to a given magnitude-limited survey.

The first observation of extra-solar lensing was made by Walsh *et al.* (1979) in their discovery of the double image of quasar QSO 0957+561A,B. The number of such multiple-imaged quasars grew rapidly. By the 1990s they had been supplemented with the discovery of Einstein rings Hewitt *et al.* (1988), giant luminous arcs in galaxy clusters (Grossman & Narayan 1988) and microlensing due to stars in the Large Magellanic Cloud (Alcock *et al.* (2000), Aubourg *et al.* (1993), Udalski *et al.* (1993)).

Gravitational lensing is now a firmly established phenomenon on a range of scales in the Universe. All instances of gravitational lensing stem from the idea that light paths respond to mass. Measurement of distortions lets us infer information about the distribution of mass in the

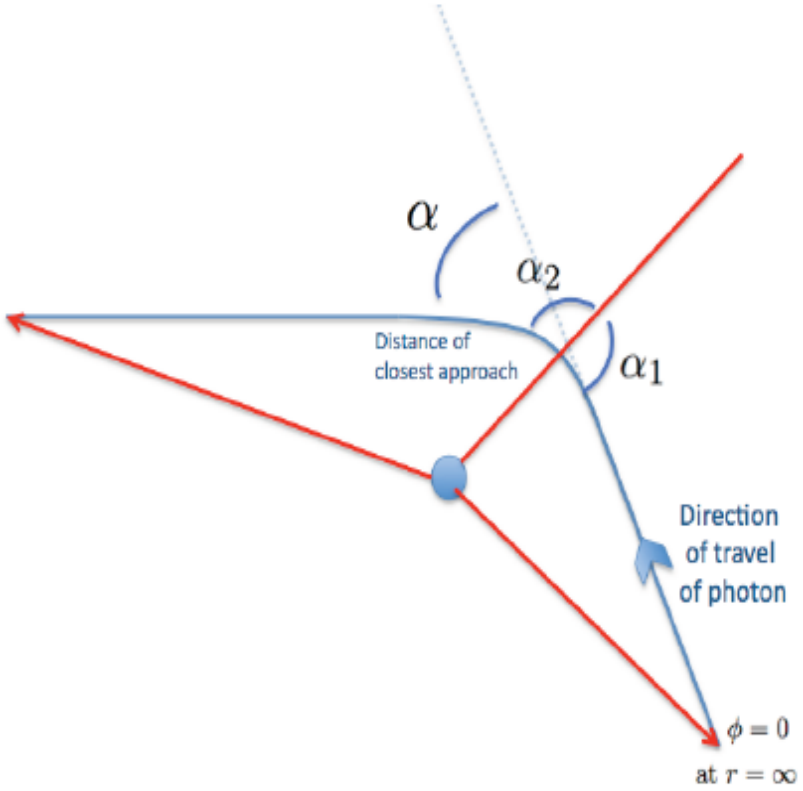

Figure 1.4: The deflection $\Delta\phi$ of the photon due to the lens can be separated into two equal parts. Anais Rassat.

Universe, whether for direct mass reconstruction or as a powerful cosmological tool. That the first practical use of gravitational lensing was as a test of gravity is an interesting point and something we will return to in this thesis.

1.7.1.2 Bend Angle

General Relativity is a theory of gravity which relates curvature of a spacetime manifold to mass present in the spacetime. Light paths lie along geodesics in the curved spacetime, producing the observed bending of light by mass. In fact, if we can make the thin lens approximation that deflection occurs over a distance much smaller than the source-lens or observer-lens distances, then we only need one equation from GR, that of the bend angle, to describe gravitational lensing.

If we take the example of a photon passing a point mass lens (as shown in Fig. 1.4), the trajectory of the photon will change by an amount $\Delta\phi$ as it passes the lens.

Writing the metric in the conformal Newtonian gauge (Mukhanov *et al.* 1992)

$$ds^2 = \left(1 + \frac{2\Psi}{c^2}\right) dt^2 - \frac{a^2}{c^2}\left(1 - \frac{2\Phi}{c^2}\right) dx^2 \tag{1.46}$$

and note that for photons following geodesics $ds = 0$ then, in the weak field limit ($|\Phi|/c^2 \ll 1$), we can rearrange as

$$dt = \frac{1}{c}\left(1 - \frac{2\Phi}{c^2}\right) dx. \tag{1.47}$$

By simple application of $time = dist./c$ we can see that the light ray propagates as if acted on by an *effective refractive index* of

$$n_{eff} = 1 - \frac{2\Phi}{c^2} = 1 + \frac{2|\Phi|}{c^2}. \tag{1.48}$$

The bend angle is then the integral along the light path of the gradient of n perpendicular to the light path (Narayan & Bartelmann 1996),

$$\hat{\alpha} = -\int \nabla_\perp n dl = \frac{2}{c^2}\int \nabla_\perp \Phi dl. \tag{1.49}$$

For the case of point mass M, the potential is of the form

$$\Phi(b, z) = -\frac{GM}{(b^2 + z^2)^{1/2}}, \tag{1.50}$$

where b is the impact parameter of the unperturbed light ray and z indicates distance along the unperturbed light ray from the point of closest approach.

Combined, these equations leave us with the bend angle equation in GR:

$$\hat{\alpha} = \frac{4MG}{bc^2}. \tag{1.51}$$

1.7.1.3 The Lens Equation

Armed with the bend angle equation from GR we look at the sketch of a typical gravitational lens system shown in Fig. 1.5 and write

$$\alpha = \theta - \beta = \frac{D_{ds}}{D_s}\hat{\alpha}, \tag{1.52}$$

where D_{ds} is the source-lens distance and D_s is the source-observer distance. Substituting in the bend angle equation produces

$$\alpha = \frac{D_{ds}}{D_s}\frac{4GM}{c^2 b}, \tag{1.53}$$

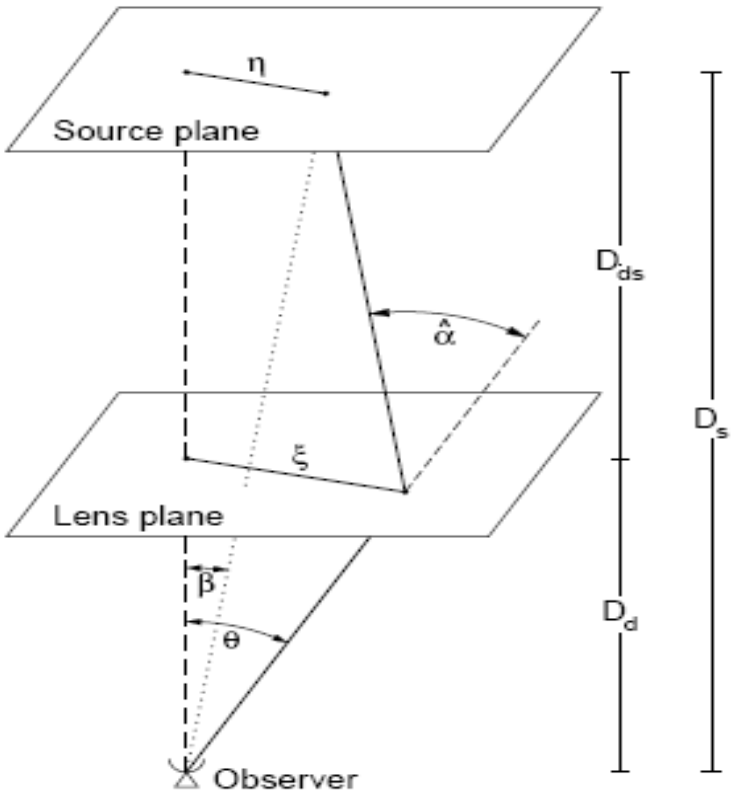

Figure 1.5: Sketch of a typical gravitational lens system showing source, lens and observer planes. Bartelmann & Schneider (2001)

.

for a point mass lens. By geometry $b = \theta D_d$, where D_d is the observer-lens distance, giving

$$\alpha = \frac{D_{ds}}{D_s D_d} \frac{4\pi GM}{c^2 \theta}. \tag{1.54}$$

This equation emphasises how important the geometry of the lensing system is to the deflection produced.

When the source lies directly on the optic axis of the lens, $\alpha = \theta$, we see a phenomenon called an Einstein Ring. The perfect circular symmetry of the system allows light to be bent at all possible deflection angles and arrive at the same observer, the image produced for a bright point source is a ring centred on the lens. The distance θ_E at which this occurs is called the Einstein radius

$$\theta_E^2 = \frac{D_{ds}}{D_s D_d} \frac{4\pi GM}{c^2}. \tag{1.55}$$

For non-point mass lenses, with surface mass density Σ a useful quantity to define is the critical

projected surface mass density,

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}}. \tag{1.56}$$

In this way, we can treat lensing by the full density perturbation field by assuming that a large number of individual small lensing effects (equivalent to point mass deflections) can be added linearly to produce an aggregate lensing effect.

1.7.1.4 Lensing Regimes

The character of gravitational lensing can be very different depending on the mass of the source, geometry of the system and hence the strength of the lensing deflection. Galaxy mass lenses, when the source galaxy is closely aligned with the optical axis, produce strong distortion to the source galaxy image. This is the *strong lensing* regime $\Sigma \geq \Sigma_{\rm crit}$. It is characterised by multiple images and the type of elongated arcs seen for example in the Abell 2218 cluster (Kneib *et al.* 1996). Time delay between multiple images with different path lengths can be a useful probe of the geometry of the Universe, giving constraints on H_0 (Paraficz & Hjorth 2010). Strong lensing sources have also been used as cosmic telescopes as they magnify sources that would otherwise be too faint to be seen by a given survey. Some of the highest redshift galaxies detected have been found via strong lensing (Kneib *et al.* 2004).

While galaxy-mass lenses are necessary to fully resolve strongly lensed images, *microlensing* has been detected in the Large Magellanic Clouds orbiting our galaxy.

As a brown dwarf or other massive body passes in front of a source star it acts as a point mass lens which can magnify the source intensity by a factor of ≥ 1.34. The relative rarity of microlensing events requires thousands of potential candidates to be monitored simultaneously. Microlensing has become an important tool in the search for MACHOs in our local Universe as well as the detection of extra-solar planets (Sackett *et al.* 2000).

The main focus of this thesis is *weak gravitational lensing*. In the regime $\Sigma \ll \Sigma_{crit}$ the distortion to the shape of lensed galaxies is very slight, certainly not obvious to the naked eye, nor do we see exotic features such as multiple images, arcs or rings. The lensing effect is still present and can be described by the convergence (magnification) and shear (two components describing acquired ellipticity). The induced shear ellipticity is of order $\sim$0.01 while galaxy intrinsic ellipticity is $\sim$0.1. Careful statistical treatment can extract the cosmic shear signal and prove a useful cosmological probe. More details are given in 1.7.3.

1.7.2 Weak Gravitational Lensing

Weak gravitational lensing deals with very small distortions to a source image. It is useful to define a scalar potential $\psi(\underline{\theta})$ which is the appropriately scaled, projected Newtonian potential of the lens or lens system (Narayan & Bartelmann 1996),

$$\psi(\underline{\theta}) = \frac{D_{ds}}{D_d D_s}\frac{2}{c^2}\int \Phi(D_d\theta, z)dz. \tag{1.57}$$

The derivative of this lensing potential with respect to theta is the bend angle as discussed above.

If a source is much smaller than the angular scale on which lens properties change, we can describe the distortion mapping by the lensing Jacobian

$$\begin{aligned} A_{ij} &= \frac{\delta\beta_i}{\delta\theta_j} \\ &= \frac{\delta}{\delta\theta_j}\left[\theta_i - \alpha_i(\theta)\right] \\ A_{ij} &= \delta^D_{ij} - \frac{\delta\alpha_i(\theta)}{\delta\theta_j} \end{aligned} \tag{1.58}$$

which can itself be written in terms of second derivatives of the lensing potential,

$$A_{ij} = \delta^D_{ij} - \frac{\delta^2\psi(\theta)}{\delta\theta_i\delta\theta_j}. \tag{1.59}$$

At this point we define the distortion matrix according to

$$A = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}. \tag{1.60}$$

where κ is the convergence, proportional to the projected gravitating mass along the line of sight. It describes the overall dilation and contraction of images. There is no photon emission or absorption involved in gravitational lensing. Combined with Liouville's theorem this tells us that surface brightness is conserved (Kristian & Sachs 1966). Therefore if gravitational lensing increases the area of an image we will see a *magnification*.

γ_1 and γ_2 are components of the 2D shear field $\gamma = \gamma_1 + i\gamma_2 = \gamma e^{i2\alpha}$. The usual definition is that γ_1 is stretching and compression along the x-axis and γ_2 is stretching and compression along the x=y axis. A source with circular isophotes would be mapped into an elliptical image. So the Jacobian describes both magnification and distortion of the source image.

It can be useful to discuss an effective convergence, κ_{eff}, defined as the convergence projected along the line of sight

$$\kappa_{eff} = \int_0^{z_d} \kappa_{\text{source},i} n(z_s) dz_s \tag{1.61}$$

where $n(z_s)$ is the distribution of source galaxies as a function of redshift.

$$\begin{aligned}\kappa_{eff} &= \int_0^{z_d} \frac{\Sigma}{\Sigma_{\text{crit}}} n(z_s) dz_s \\ &= \Sigma \int_0^{z_d} \frac{4\pi G}{c^2} \frac{D_d D_{ds}}{D_s} n(z_s) dz_s \\ &= \Sigma \frac{4\pi G}{c^2} D_d \int_0^{\chi_d} \frac{D_{ds}}{D_s} n(\chi_s) d\chi_s. \end{aligned} \tag{1.62}$$

This will become very useful for cosmic shear tomography, where we cut a galaxy distribution into redshift slices.

Generalising to multipe source planes we can write (Munshi *et al.* 2006)

$$\kappa = W(\chi_d) \int (\delta(\chi)) d\chi_d \tag{1.63}$$

where

$$W(\chi) = \frac{3}{2} \frac{\Omega_m H_0^2}{c^2} (1 + z(\chi)) \int n(\chi_s) \frac{(\chi_s - \chi)\chi}{\chi_s} d\chi_s \tag{1.64}$$

is the lensing weight function or lensing efficiency function.

In Fourier space

$$\tilde{\gamma}(l) = e^{i2\beta} \tilde{\kappa}(l) \tag{1.65}$$

so if we can measure a shear field on the sky we can relate it to the convergence and map the projected gravitational mass map of the region.

1.7.3 Cosmic Shear

Cosmic shear is a term we will use to describe the use of the very small distortions caused by weak gravitational lensing to learn about the Universe, particularly cosmological parameters. The idea of using cosmic shear to learn about the growth of structure with redshift and the expansion history of the Universe can be traced back to a lecture given by Richard Feynman at Caltech in 1964 (Refregier 2003). The propagation of light was then studied by theorists ((Kristian & Sachs 1966),(Schneider & Weiss 1986),(Lee & Paczynski 1990)) and the necessary statistics for comological use were computed and discussed ((Blandford *et al.* 1991), (Kaiser 1992), (Miralda-Escude 1991)).

At very nearly the same time, four groups ((van Waerbeke *et al.* 2000), (Kaiser *et al.* 2000), (Bacon *et al.* 2000), (Wittman *et al.* 2000)) produced measurements of weak gravitational lensing from statistical treatment of distant galaxies.

We have already introduced the 2pt-correlation function, and its Fourier transform, the Power Spectrum, in the context of the matter distribution and galaxies. Cosmic shear is a very promising cosmological tool partly because it offers an unbiased tracer of the underlying dark matter distribution

We can show that the shear power spectrum is exactly equivalent to the convergence power spectrum. We are free to work with whichever quantity is most convenient. We will call the common statistic the lensing power spectrum for ease.

It is often convenient to measure shear and ellipticity, not in terms of their real and imaginary parts, but as tangential and cross components:

$$\epsilon_+ \equiv -Re(\epsilon e^{-2i\phi});\ \epsilon_\times \equiv -Im(\epsilon e^{-2i\phi}) \tag{1.66}$$

where ϕ is the angle to the line of sight, ϵ is galaxy ellipticity and Re/Im take the real and imaginary parts of an expression respectively. $+$ denotes a component tangential to the line of sight and $\times$ denotes the component at 45^o to the line of sight. Similar expressions can be defined for shear and observed ellipticity.

With these definitions we can define a very practical measure of the cosmic shear signal, the 2-pt. correlation functions

$$\xi_\pm(\theta) \equiv < \gamma_+\gamma_+ > \pm < \gamma_\times\gamma_\times > . \tag{1.67}$$

This type of function, the correlation function ξ, is very common and takes the product of two quantities (in this case a pair of galaxies) and averages them over independent realisations of the system in question. Now in practice we do not have access to multiple realisations of the Universe. Instead we invoke an assumption of ergodicity, that is we assume that averaging over a sufficiently large volume is equivalent to an ensemble average (Pan & Zhang 2010). In general a correlation function is often calculated as a function of separation. In this case the average is over pairs of galaxies separated by some fixed distance.

We will derive the lensing power spectrum by expanding the effective convergence in spherical harmonic components. This is a derivation I produced as part of my PhD work. The derivation in spherical harmonics is exact, we will then apply the Limber approximation to find the more computationally tractable approximate form which is used throughout the rest of this thesis.

Defining the spherical harmonic transform

$$\kappa_{lm} = \int d\Omega Y^*_{lm}(\theta)\kappa(\theta), \tag{1.68}$$

and the inverse:

$$\kappa(\theta) = \sum_{l,m} \kappa_{lm} Y_{lm}(\theta), \tag{1.69}$$

we can expand

$$\kappa(\theta) = \Sigma_{l,m} \int d\Omega' Y^*_{lm}(\theta')\kappa(\theta')Y_{lm}(\theta) \tag{1.70}$$

$$= \int d\Omega' \kappa(\theta') \sum_{l,m} Y^*_{lm}(\theta')Y_{lm}(\theta) \tag{1.71}$$

and write

$$\kappa_{lm} = \int d\Omega Y^*_{lm}(\theta) \int W(\chi) \frac{1}{V} \sum_k \tilde{\delta}(k) e^{ik.x} d\chi \tag{1.72}$$

where we have used the definition of the FT f $\delta(k)$.

Then, using the relation $e^{ik.x} = 4\pi \sum_{l,m} i_l j_l(k\chi) Y^*_{lm}(\hat{k}) Y_{lm}(\hat{\Omega})$,

$$\kappa_{lm} = \int d\chi W(\chi) \frac{1}{V} \sum_k \tilde{\delta}(k) 4\pi \sum_{l',m'} i^{l'} j_{l'}(k\chi) Y_{l'm'}(\hat{k}) \int d\Omega Y^*_{lm}(\theta) Y_{l'm'}(\theta) \tag{1.73}$$

and from orthogonality $\int d\Omega Y^*_{lm}(\theta) Y_{l'm'}(\theta) = \delta_{Kll'}\delta_{Kmm'}$,

$$\kappa_{lm} = \int d\chi W(\chi) \frac{1}{V} \sum_k \tilde{\delta}(k) 4\pi i^l j_l(k\chi) Y_{lm}(\hat{k}) \tag{1.74}$$

Now define the power spectrum as

$$C_l = \left\langle \left\langle K^{lm}_* K^{lm} \right\rangle_{sky} \right\rangle_m \tag{1.75}$$

and consider

$$\langle \kappa^*_{l'm'} \kappa_{lm} \rangle_{sky} = \left\langle \int d\chi' W(\chi') \frac{1}{V'} \sum_{\underline{k}'} \tilde{\delta}^*(\underline{k}') 4\pi (i^{l'})^* j_{l'}(k'\chi') Y^*_{l'm'}(\hat{\underline{k}}') \int d\chi W(\chi) \frac{1}{V} \sum_{\underline{k}} \tilde{\delta}^*(\underline{k}) 4\pi (i^l)^* j_l(k\chi) Y^*_{lm}(\hat{\underline{k}}) \right\rangle \tag{1.76}$$

$$= \int d\chi' W(\chi') d\chi W(\chi) \frac{1}{V'} \sum_{\underline{k}} P(k) (4\pi)^2 (i^{l'})^* i_l j_{l'}(k'\chi') j_l(k\chi) Y^*_{l'm'}(\hat{\underline{k}}') Y^*_{lm}(\hat{\underline{k}}) \tag{1.77}$$

For the lensing power spectrum we are only interested in the case $l' = l, m' = m$, averaging over m and using $< Y_{lm}(\underline{\theta} Y^*_{lm}(\underline{\theta})) >_m = \frac{1}{4\pi}$:

$$\ll \kappa^*_{lm} \kappa_{lm} >_{sky} >_m = \int d\chi' W(\chi') \int d\chi W(\chi) \frac{1}{V'} \sum_{\underline{k}} P(k) (4\pi)^2 j_l(k\chi') j_l(k\chi) \frac{1}{4\pi} \tag{1.78}$$

$$C_l = \int d\chi' W(\chi') \int d\chi W(\chi) \frac{4\pi}{V'} \sum_{\underline{k}} P_\delta(k) j_l(k\chi') j_l(k\chi) \tag{1.79}$$

and redefining the Fourier Transform over an infinite volume we get

$$C_l = \int d\chi' W(\chi') \int d\chi W(\chi) \int \frac{2k^2}{\pi} dk P_\delta(k) j_l(k\chi') j_l(k\chi). \tag{1.80}$$

This is the exact lensing integral but it is rather cumbersome to integrate. We can proceed without significant loss of accuracy in the rest of this thesis if we make the common approximation that the area on the sky we're interested in is small enough that we deal with large enough l to allow us to approximate

$$j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+1/2}(x) \tag{1.81}$$

$$\lim_{l\to\infty} j_l(x) = \sqrt{\frac{\pi}{2l+1}} \delta_D(l+1/2-x) \tag{1.82}$$

this is known as the Limber approximation and gives us

$$C_l = \int d\chi' W(\chi') \int d\chi W(\chi) \int \frac{2k^2}{2l+1} dk \frac{1}{\chi} \delta_D(\frac{l+1/2}{\chi} - k) \delta_D(l+1/2-k\chi') P_\delta(k). \tag{1.83}$$

Integrating with respect to k produces

$$C_l = \int \int d\chi' d\chi W(\chi') W(\chi) \frac{1}{\chi^2} \delta_D(\chi-\chi') P_\delta(\frac{l+1/2}{\chi}). \tag{1.84}$$

We can integrate over χ' and replace all the χ' with χ, giving the simplified version of the projected angular power spectrum

$$C_l = \int \frac{d\chi}{\chi^2} W^2(\chi) P_\delta\left(\frac{l}{\chi}\right) \tag{1.85}$$

where we have made the additional approximation $l + 1/2 \equiv l$.

Cosmic shear surveys acquire distance information as well as measurements of galaxy shapes. Sometimes the information is in the form of high-quality spectroscopic redshifts that use the doppler shift of galaxy spectra to calculate redshift. This is a costly and time consuming procedure. Because of the large number of source galaxies in a cosmic shear survey the redshift information is more likely to be from photometric measurements. Here the intensity of a source in different frequency bands is compared to a set of templates calibrated against spectroscopic measurements and a redshift determined. The photometric approach is much quicker but less accurate than spectroscopy though new techniques and improvements are an active area of research. Machine learning on a training set of known redshift galaxies is greatly improving the accuracy of photometry.

The most common way to use this redshift information is cosmic shear *tomography*. Instead of projecting the entire galaxy sample onto the sky and measuring the 2D correlation function, tomgoraphy cuts the galaxy distribution into slices in redhsift and calculates the correlation function

on each slice (auto-correlations) and between all the combinations of slices (cross-correlations). We can write the shear angular power spectrum between two redshift bins i and j as

$$C_{ij}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) P_\delta(\frac{l}{\chi}). \tag{1.86}$$

More distant shells have stronger lensing signals as the light from those sources passes more mass on the way to the observer but we must take into account that the z-shells are correlated with each other- the mass in lower-z shells is lensing light from higher-z sources. It is important that we calculate the full covariance matrix between shells when evaluating likelihoods.

Tomography is a useful extension to lensing but it is important to consider the goal when deciding on a survey and analysis strategy. Tomography does not greatly improve estimation of the matter power spectrum but it is very useful when, for example, attempting to constrain the evolution of the equation of state of dark energy, $w(a)$.

The use of redshift information also introduces a number of new systematic effects, most importantly redshift accuracy and the fraction of redshifts which are catastrophically mis-estimated. Poor redshift accuracy can lead us to think galaxies are in a particular z-shell when they are in reality in another. They also limit the utility of increasing the number of tomographic slices to arbitrarily large values.

An alternative technique for using available redshift information is known as 3D weak lensing (Heavens 2003). This does away with redshift slicing or any 2D statistics and instead treats each galaxy as a noisy estimator of the full 3D shear field. It is most convenient to write the 3D power spectrum in terms of spherical harmonics and a bessel function expansion

$$\pm g_{lm}(k) = \sqrt{\frac{2}{\pi}} \int d^3\underline{\chi} n(\underline{\chi}) \gamma_\pm(\underline{r}) j_l(k\underline{\chi}) {}_{\pm 2}Y_l^m(\hat{\underline{n}}). \tag{1.87}$$

This approach avoids the loss of information from binning galaxies but with the downside of vastly increased computation time, making it, at present, impractical for the large datasets that current and future surveys are set to produce.

Having dealt in detail with 2-point statistics of cosmic shear we can generalise the formalism into 3-point statistics, writing the convergence 3-pt. correlation function $< \tilde{\kappa}_i(\underline{l}_1)\tilde{\kappa}_j(\underline{l}_2)\tilde{\kappa}_k(\underline{l}_3) >$ and defining its Fourier transform, the bispectrum, B_{ijk}

$$< \tilde{\kappa}_i(\underline{l}_1)\tilde{\kappa}_j(\underline{l}_2)\tilde{\kappa}_k(\underline{l}_3) > \equiv (2\pi)^2 B_{ijk}(\underline{l}_1, \underline{l}_2, \underline{l}_3)\delta_D(\underline{l}_{123}). \tag{1.88}$$

$\delta_D(\underline{l}_{123})$ ensures that the three vectors $\underline{l}_1, \underline{l}_2, \underline{l}_3$, form a traingle in Fourier space. The measured bispectrum will depend on the geometry of the triangle configurations used. The bispectrum is

the 3-pt. analogue of the power spectrum. The power spectrum is only sensitive to Gaussian perturbations while 3-pt. statistics are the lowest order necessary to describe non-Gaussianity. There are two sources of non-Gaussianity in the cosmic shear field:

- *Primoridial Non-Gaussianity* imprinted in the early Universe. This is a prediction of slow-roll inflation and some other inflation models as well as models of reheating which allow a spatial dependence. The amount of non-Gaussianity is often parameterised by the quantity f_{NL} via $\Phi(\underline{(x)}) = \phi(x) + f_{Nl}[\phi^2(x) - < \phi^2(x) >]$.

- *Non-Gaussianity due to non-linear galaxy clustering.* As small background perturbations in the gravitational potential (and hence the mass distribution) grow under the influence of gravity highly non-linear regions ($\delta \geq 1$) form on small scales. Non-linear clustering greatly enhances the amplitude of the mass fluctuation and the lensing signal on scales below 1°. This process will induce non-Gaussianity in the shear field even if the primordial potential was Gaussian.

A potentially interesting property of the shear bispectrum is that it is unaffected by the intrinsic ellipticity distribution of galaxies if that distribution is symmetric with a mean of zero.

1.7.4 Shear Measurement

An optical cosmic shear survey captures images of small, faint galaxies. It is far from straightforward to convert these images into a measure of the cosmic shear field. After the light is captured, a number of observational distortions must be corrected for. These include CCD pixelisation, astrometric calibration and most importantly the effect of the point spread function (PSF) which distorts a circular image due to the telescope optics themselves, we discuss these effects in more detail in chapter 6.

Even if these effects can be dealt with to sufficient accuracy, measurement of the ellipticity of a source galaxy and the extraction of the cosmic shear signal remains difficult. The induced shear on any galaxy is a very small effect ($\sim$1%). Moreover, galaxies are intrinsically elliptical so it is impossible to disentangle the shear effect from the intrinsic ellipticity for a single galaxy image.

One popular method for extracting shear from galaxy images was the Kaiser, Squiers & Broadhurst, or KSB, method (Kaiser *et al.* 1995). This used the simple approach of calculating the

quadrupole moments of the surface brightness distribution I($\underline{x}$) for each source galaxy:

$$Q_{ij} = \int d^2x x_i x_j w(x) I(\underline{x}), \tag{1.89}$$

where $w(x)$ is a weight function, usually assumed to be Gaussian. This itself is a very noisy estimator as most galaxies in a particular sample will be faint.

If measured accurately the quadrupole moments can be used to reconstruct the observed galaxy ellipticities (Bridle *et al.* 2009):

$$\epsilon^{obs} \equiv \epsilon_1^{obs} + i\epsilon_2^{obs} = \frac{Q_{xx} - Q_{yy} + 2iQ_{xy}}{Q_{xx} + Q_{yy} + 2(Q_{xx}Q_{yy} - Q_{xy}^2)^{1/2}}, \tag{1.90}$$

where we introduce the complex notation $\epsilon = \epsilon_1 + i\epsilon_2$, $\gamma = \gamma_1 + i\gamma_2$. For a simple galaxy that has concentric, elliptical isophotes (contours of constant brightness) with major axis a, minor axis b and angle θ between the positive x axis and the major axis,

$$\epsilon_1 = \frac{a-b}{a+b} cos(2\theta), \tag{1.91}$$

$$\epsilon_2 = \frac{a-b}{a+b} sin(2\theta). \tag{1.92}$$

Observed galaxy ellipticites are related to the shear via

$$\epsilon^{obs} = \frac{\epsilon^{int} + g}{1 + g^* \epsilon^{int}} \tag{1.93}$$

where ϵ^{int} is the galaxy intrinsic ellipticity and g is the reduced shear $g \equiv \frac{\gamma}{1-\kappa}$. For sufficiently weak lensing we can approximate $\epsilon^{obs} = \epsilon^{int} + \gamma$.

When observed ellipticities have been measured we designate a patch of the sky over which we assume shear is constant because light from all the sources has passed the same LSS en route to the observer. Averaging over the observed ellipticities on this patch leaves the shear signal $\langle \epsilon^{obs} \rangle = \gamma$, if we assume that the intrinsic ellipticities of galaxies are randomly distributed on the sky and average to zero $\langle \epsilon^{int} \rangle = 0$. This assumption proves to be inaccurate. The physics of intrinsic galaxy alignments becomes a focus of this thesis from chapter 2.

There have been many improvements in cosmic shear measurement since the KSB method was pioneered. Often shear measurement pipelines fit distortions to a circular galaxy to observed ellipticities or decompose images into summations of "shapelet" basis functions. For more detail on shear measurement, the latest techniques and accuracy tests see chapter 6.

There is a useful decomposition of the shear field that is used to check for the effects of systematics on the lensing signal. We have already seen that the shear field is directly related to the projected mass distribution. In fact both are simply different 2$^{\text{nd}}$ derivatives of the gravitational

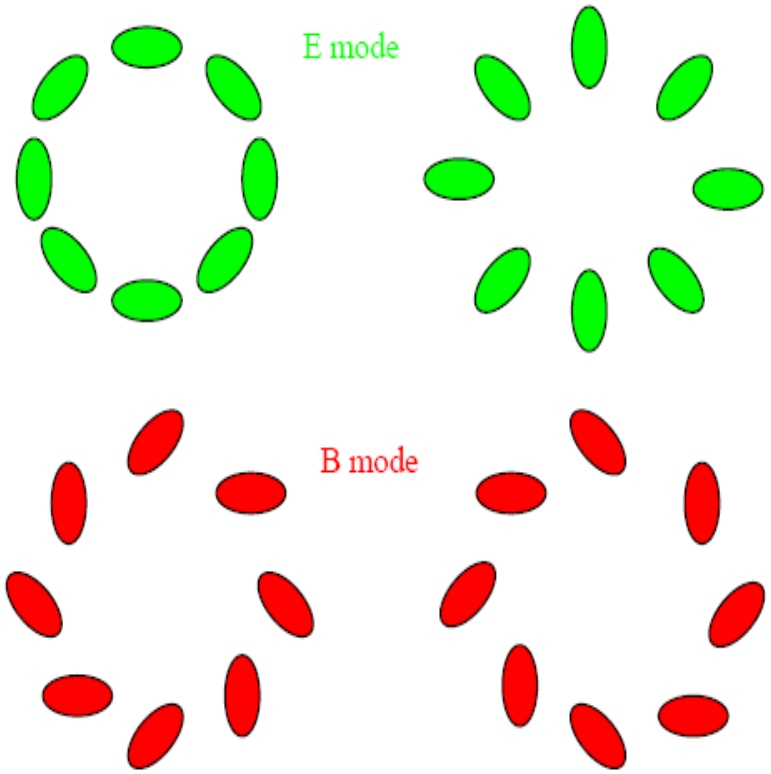

Figure 1.6: Illustrative E and B modes: the E modes show what is expected around overdensities (left) and underdensities (right). The B mode patterns should not be seen . Munshi *et al.* (2006)

potential projected along the line of sight. Mass concentrations produces certain circular effects in the shear field known as E-modes. Fig. 1.6 shows the tangential E-mode associated with a mass overdensity which looks qualitatively similar to the strong-lensed arcs seen around clusters. Also shown is the complementary radial E-mode oberved (on much larger scales) around underdensities. This effect is less apparent because of the smaller contrast between a void and the average mass background.

It is possible to extract a 2nd scalar quantity from a general complex field, this has a curl-like behaviour and is called the B-mode. The shear field is real and the distortions produced by mass cannot produce B-modes (some particularly complex massive structures can produce a very small B-mode signal). In contrast, many systematic effects which contaminate the cosmic shear signal, including the optics' PSF and systematic changes to source depth, produce equal quantities of E- and B-modes. Measuring a B-mode signal from a survey which is consistent with zero is a good check on the removal of these systematics.

The E/B nomenclature comes from analogy with polarisation fields, which in turn comes from the language of electrodynamics where E denotes the electric field and B the magnetic field whose divergence is zero. We can write E and B correlation functions in terms of the standard correlation functions

$$\xi_E(\theta) = \frac{1}{2}\left[\xi_-(\theta) + \xi'_+(\theta)\right] \tag{1.94}$$

$$\xi_B(\theta) = -\frac{1}{2}\left[\xi_-(\theta) - \xi'_+(\theta)\right] \tag{1.95}$$

$$\text{where } \xi_+(\theta) + 4\int_0^\theta \frac{d\vartheta}{\vartheta}\xi_+(\vartheta) - 12\theta^2\int_0^\theta \frac{d\vartheta}{\vartheta^3}\xi_+(\vartheta) \tag{1.96}$$

which avoids the need to know the correlation functions at large scales. Even with this forumlation, the E/B-mode decomposition will be ambiguous without a full-sky survey (Munshi *et al.* 2006).

An interesting way to think about the restrictions on properties of the shear field is that there are six free parameters which describe locally linear transformations: translation ($\times 3$), rotation, dilation and shear ($\times 2$). Even after we discard translation as not readily observable, shear and convergence can be fully specified by three of the remaining four degrees of freedom. This restriction helps explain the lack of B-modes from cosmic shear and can be seen in the symmetry of the 2×2 shear distortion matrix.

1.7.5 WGL Systematics

There are a number of important systematics that must be accounted for in any practical application of WGL as a cosmological tool. We briefly summarise a few here. The most important cosmic shear systematic comes from the intrinsic alignment of galaxy ellipticities. This is dealt with thoroughly throughout the thesis. For more information on shear measurement in the presence of systematics see chapter 6.

The *point-spread function* (PSF) describes the distortion of a point source due to features of the telescope optics. Particularly strong on ground-based surveys due to atmospheric effects, it is still very much present for space-based missions. The distortion of galaxy shapes must be tightly controlled for cosmic shear surveys. This is usually done by calibration with star images which act as point sources (Paulin-Henriksson *et al.* 2008).

When galaxies are very close on the sky *overlapping isophotes* can produce "peanut shaped" images and suggest an alignment of ellipticities where none exists.

CCD Effects, properties of the CCD chips which collect the light in surveys can themselves act as systematic effects. Nonlinear response (Van Waerbeke *et al.* 2006), particularly if it varies

from pixel to pixel, can induce bias in the measured shear. Charge Transfer Inefficiency (CTI) in the way electrons are read out of the CCD pixels can add ellipticity to galaxy shapes, biasing our shear measurements (Massey *et al.* 2010).

As with all cosmological probes, certain *theoretical uncertainties* limit our ability to measure shear to the accuracy we would desire. Prominent areas of ignorance include the details of galaxy formation and evolution for different classes of galaxy, the nonlinear structure formation at small scales, galaxy biasing and the cross-correlation of different probes on the same patch of the sky. More powerful analytic tools and larger simulations with more particles are complementary in answering some of these open questions.

The sheer number of galaxies which will be measured by future cosmic shear surveys creates its own problems. Ideally we would record accurate spectroscopic redshifts for each galaxy to aid our tomographic analysis. In practice this proves too expensive and time-consuming. Instead we rely on photometric redshifts which are less accurate by easier and cheaper to gather (Hildebrandt *et al.* 2010). See chapters 4 & 5 for further discussion of this point.

1.7.6 Cosmic Shear Surveys

Cosmic shear was first detected at the millenium by four independent groups (Kaiser et al. 2000; Refregier et al 2000; van Waerbeke et al. 2000; Wittman et al. 2000). A number of surveys were carried out to turn this detection into a measurement of cosmological parameters, which are reviewed in Refregier (2003).

The most recent surveys are the Red-sequence Cluster Survey (RCS) (45) of 53 square degrees, VIRMOS (136) measurements of 8.5 square degrees, the CTIO shear survey (65) relatively shallow measurements of 70 square degrees, the Garching Bonn Deep Survey (GaBoDS) (38) of 13 square degrees, the Hubble Space Telescope COSMOS survey (89) of 2 square degrees and the Canada France Hawaii Telescope Legacy Survey (CFHTLS) of order 100 square degrees (analysis of 57 square degrees is given in 31).

The CFHTLS team is making their final checks before publishing their analysis on the full dataset. Pan-STARRS has already started taking first data. Meanwhile several surveys are due for first light in the coming year: the KIlo-Degree Survey (KIDS) has been long awaited on the Very Large Telescope (VLT) Survey Telescope (VST) and will survey 2000 square degrees, the Subaru Measurement of Images and Redshifts (SuMIRe) will use the new HyperSuprimeCam instrument and the Dark Energy Survey (DES) is also installing a new wide-field corrector.

Further into the future the Large Synoptic Survey Telescope (LSST) has already cast the com-

bined primary/tertiary lens and is sufficiently ambitious to be classed as a "Stage IV" experiment (Albrecht et al 2007). The proposed space project Euclid is being assessed for final down-selection this summer and the US has a project called WFIRST.

It has been identified that the radio may offer the most exciting possibilities for measuring cosmic shear in the far future. The point spread function is well known because it is fixed by the positions of the radio antennae, and spectral information may also be obtained overcoming the photometric redshift problem. The ultra-ambitious radio project the Square Kilometer Array (SKA) has a number of ongoing pathfinders.

1.7.7 Beyond Cosmic Shear

We have already noted the ability of WGL surveys to reconstruct projected mass maps from measurements of the convergence, κ. Beyond Cosmic Shear, which we reserve for large-scale shear surveys aimed at the measurement of cosmological parameters, weak lensing can be a useful tool in understanding massive structures.

Galaxy Clusters have been a useful complement to standard cosmic shear work. Clusters trace the highest peaks in the density field. Studies of cluster number counts as a function of mass and redshift can act as probes of cosmology. There are difficulties taking into account projection effects but cosmic shear allows a direct reconstruction of the mass field (Sheldon *et al.* 2009) and access to peak statistics.

Gravitational lensing has proven to be a tool well-suited to probing the *properties of DM halos*. Virialised halos seem to have a near universal profile with radius- the Navarro-Frenk-White (NFW) Profile over a wide range of halo mass. Inner regions of halos can be probed by strong lensing in combination with galaxy dynamics (Koopmans *et al.* 2006) while the dark matter dominated outer regions are amenable to WGL (Hoekstra *et al.* 2002).

Flexion

Cosmic shear and convergence are the 2^{nd} angular derivatives of the gravitational potential, projected along the line of sight. If we consider the next order of differentiation we find two independent combinations of the 3^{rd} derivatives which are termed "flexion."

The flexion effects arise when the projected mass distribution of a lens has a spatial gradient steep enough to change the induced shear from one side of a source galaxy to the other. The variation in shear across the images produces two distinct distortions:

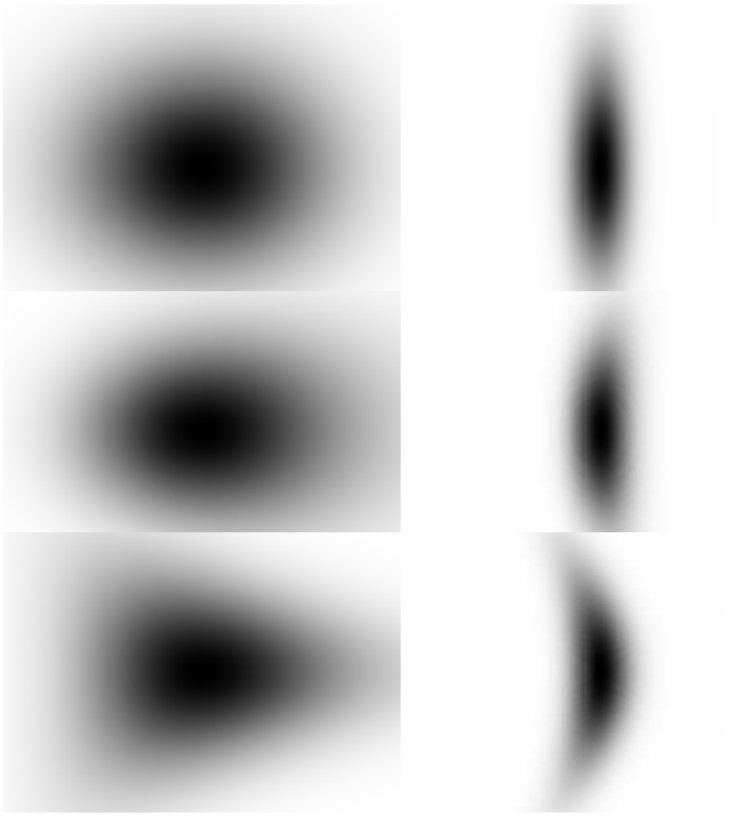

Figure 1.7: Effect of 1-flexion and 3-flexion on circular and elliptical (e=0.9) Gaussian sources. Top panels: Unlensed objects; Middle panel: $F_1 = 0.2\text{arcsec}^{-1}$; Bottom Panel: $G_1 = 0.7\text{arcsec}^{-1}$. Bacon & Schäfer (2009)

- 1-Flexion, a vector distortion which skews the image shape.
- 3-Flexion, a spin-3 distortion which changes circular images into trefoils.

Figure 1.7 illustrates the curvature effects from the 1- and 3-flexions.

The amplitude of the flexion effect is much smaller than that of lensing but the intrinsic curvature of galaxies is much less than their intrinsic ellipticities, helping the achievable signal-to-noise. Flexion has been found to be particularly useful in the reconstruction of mass maps around galaxy clusters. In these cases the light deflection is often too small for strong lensing effects to manifest and the areas of interest are too small to provide enough source galaxies for a useful cosmic shear analysis. Flexion is a useful tool to fill the gap between the regimes. Flexion has been detected in current space-based data (Velander *et al.* 2011) and is expected to become an inreasingly important probe as future surveys with more statistical power become available (Munshi *et al.* 2011).

1.7.8 Beyond the Optical

The main work of this thesis will focus on weak gravitational lensing at optical wavelengths of distant galaxy images but it is worth noting that there are potentially useful WGL effects at other wavebands and from other sources.

Galaxies studied by *radio and near IR* surveys will also have an associated cosmic shear signal. At present, the number density of sources resolved by radio surveys is much smaller than equivalent optical datasets but some preliminary work has been done which has detected lensing E-modes in radio data (Chang *et al.* 2004). As more ambitious radio facilities such as the Square Kilometre Array (SKA) become a reality, the number density will increase by orders of magnitude, making cosmic shear analysis a real possibility (Blake *et al.* 2004). Recent work by Brown & Battye (2011*b*), Brown & Battye (2011*a*) have suggested that the polarisation of galaxy light measured by redshift surveys may provide us with information on the intrinsic ellipticity of source galaxies. This may provide an effective route for reducing the impact of intrinsic alignment systematics on the cosmic shear signal. A lack of correlation between optical and radio intrinsic ellipticities has also been noted (Patel *et al.* 2010), improving the properties of optical-radio lensing cross-correlations.

The *Cosmic Microwave Background* will be weakly lensed by the LSS in the Universe. Lensing broadens the CMB acoustic peaks and enhances power at small scales. CMB-lensing can probe information at very large scales and high redshifts. Lensing of the polarisation signal produces B-modes from pure E-modes and can act as a contaminant when searching for primordial B-modes generated during inflation. WGL induces non-gaussianity into the CMB signal. This can allow us to reconstruct the lensing potential from the CMB sky, providing information about the mass distribution responsible or as a prelude to subtracted lensing contamination from CMB measurements (Lewis & Challinor 2006).

Much like the CMB, emission of *21cm radiation* from neutral gas prior to recombination is a diffuse background source subject to gravitational lensing (Metcalf 2007). The lensing effect is weaker than in the CMB and unlike the CMB there is considerable power at small angular scales. Lensing also makes the 21cm distribution non-Gaussian, combined with tomographic information this can be used for lensing reconstruction. If made practicable, 21cm lensing could allow us to access redshifts of 15-30, bridging the gap between galaxy lensing surveys and the CMB.

1.8 Thesis Structure

In general I assume the geometry of the Universe is flat throughout, with Gaussian and adiabatic primordial fluctuations and no running of the spectral index, $n_a = 0$. Throughout I assume it is valid to make the Limber approximation when calculating the projected angular power spectrum. Particular assumptions and fiducial models are given within each chapter.

Chapter Overview

- Chapter 2: Demonstration of biasing in measurements of cosmological parameters in the presence of galaxy Intrinsic Alignment. Parameterised IA models- including Halo Model for the first time. Complementary use of shear-position data. Joint constraints on IA parameters, marginalisation and unbiased constraints on cosmology.

- Chapter 3:The use of cosmic shear and galaxy redshift information to probe deviations from General Relativity. Degeneracy IAs/Modified Gravity. The mitigating effect of using galaxy redshift information inherent to any cosmic shear survey. Introduction of new fitting formulae for 2-parameter MG model.

- Chapter 4: Application of the MG/IA formalisms presented in the previous chapter to future cosmic shear surveys. Effect of survey parameters. Effect of photometric redshift information. Possible benefits of spectroscopic redshift information in a realistic survey framework.

- Chapter 5: Initial design study for DESpec, a proposed spectroscopic follow-up to a photometric cosmic shear survey. Analysis of the impact of spec-z information on WGL, Redshift Space Distortions (RSDs) as a cosmological probe. Joint constraints from WGL & RSD for DE and MG.

- Chapter 6: Describes work on the Great 08 shear measurement challenge for cosmologists and computer scientists. The difficulty of measuring cosmic shear from real images, the concept of the challenge, production of simulations and results. Testing of data management architecture and shear simulation for Dark Energy Survey (DES), a stage-III cosmic shear survey seeing first light in 2011.

- Chapter 7: Concludes the work in this thesis and sketches some directions for future investigation.

CHAPTER 2

SIMULTANEOUS CONSTRAINTS ON INTRINSIC ALIGNMENTS AND COSMOLOGY FROM CURRENT DATA

2.1 Introduction

Weak gravitational lensing is one of the most promising observational tools available to cosmologists for studying the recent accelerated expansion of the Universe. Images of distant galaxies appear distorted due to the bending of light as it passes through the gravitational potential of the intervening matter. The image distortion can be described as a shearing of the original galaxy shape. This "cosmic shear" can be exploited to study the matter distribution in the Universe and the growth of structure. Cosmic shear has the potential to be one of the most powerful probes both of dark matter and of dark energy (Peacock & Schneider 2006; Albrecht *et al.* 2006) which together are thought to make up about 95% of the energy budget of the Universe (e.g. Spergel *et al.* 2007).

For any individual galaxy it is impossible to separate the small cosmic shear distortion from the intrinsic ellipticity of the galaxy shape. However, the light from physically close galaxies will follow a similar trajectory to the observer, passing through spacetime curved by the same gravitational fields. Therefore these galaxies will acquire the same cosmic shear distortion. If it is assumed that galaxy shape and orientation are randomly assigned across the sky then the cosmic shear signal can be retrieved by averaging the ellipticities of a number of galaxies close on the sky.

In practice this assumption of randomly distributed galaxy shapes is unrealistic. Spatially localised galaxies are expected to have formed within the same large-scale gravitational field, which is likely to cause an alignment in their intrinsic ellipticities (Heavens & Peacock 1988; Catelan *et al.* 2001; Jing 2002; Aubert *et al.* 2004). See Schäfer (2008) for a recent review. This intrinsic alignment (IA) effect will appear as a systematic error in estimates of cosmological parameters extracted from cosmic shear data, unless accurately taken into account.

Two types of IA affect the measured cosmic shear signal. Physically close galaxies form in the same large-scale gravitational potential so share a preferred ellipticity orientation. This is the Intrinsic-Intrinsic (II) correlation. Background galaxies will be lensed by foreground gravitational potentials which govern the orientation of foreground galaxies. This causes an anti-correlation between the foreground/background galaxies known as the Gravitational-Intrinsic (GI) correlation.

Broadly, there are two approaches to dealing with IAs. The first, known as "nulling", places steps in the weak lensing pipeline which remove the IA signal. The effect of II correlations can be removed by downweighting physically close galaxy pairs (Heymans *et al.* 2004*a*; King & Schneider 2003; Heymans & Heavens 2003; Takada & White 2004). The GI term is more problematic as it affects all pairs of galaxies which are not physically close. In principle a particular linear combination of tomographic shear power spectra can be used to remove the GI signal if the redshifts of the galaxies are known (Joachimi & Schneider 2008, 2009). The other approach tries to model the IA signal we would expect for a particular survey. These IA contributions can then be incorporated into the predictions for the measured shear signal and any free parameters in the IA model can be varied and marginalised over (Albrecht *et al.* 2006; Bridle & King 2007; Bernstein 2009; Joachimi & Bridle 2009). This is the approach we investigate. To carry this out it is necessary to make physically motivated models for IAs. Ignoring IAs completely will bias estimates of cosmological parameters (Hirata & Seljak 2004; Mandelbaum *et al.* 2006*c*; Hirata *et al.* 2007; Bridle & King 2007). In addition, other cosmological data sets, such as galaxy surveys, can be used to gain empirical knowledge about the IA effect, as well as its variation with luminosity, colour and any other variables.

Some of the modelling approaches make simultaneous use of the galaxy shear - galaxy shear correlation function and the galaxy position - galaxy shear correlation function, as measured from the same imaging survey (Albrecht *et al.* 2006; Bernstein 2009; Joachimi & Bridle 2009), where "shear" has contributions from both gravitational and intrinsic shear. This was first identified as an important additional statistic for learning about galaxy formation by Hu & Jain (2004) and was recently suggested as a powerful tool for removing IAs by Zhang (2008), which was demonstrated

in Joachimi & Bridle (2009).

Cosmic shear was first detected observationally a decade ago (Bacon *et al.* 2000; Van Waerbeke *et al.* 2000; Wittman *et al.* 2000; Kaiser *et al.* 2000) and the strength of the signal as a function of galaxy separation has since been measured by many teams, most recently by Fu *et al.* (2008) and Massey *et al.* (2007*a*) (see also Schrabback *et al.* 2009). A compilation of recent cosmic shear results including a homogeneous additional treatment of systematics was made public by Benjamin *et al.* (2007). The use of photometric redshift information is now starting to allow the signal to be evaluated as a function of galaxy redshift (Massey *et al.* 2007; Schrabback *et al.* 2009). The cosmological constraints in these recent papers are calculated assuming there are no IAs.

However, if IAs are non-negligible then ignoring them will lead to a bias on cosmological parameters. The bias in σ_8 was estimated to be between 1 and 20 per cent for a CFHTLS-like survey by Mandelbaum *et al.* (2006*a*) and Hirata *et al.* (2007). Schneider & Bridle (2010*a*) and Mandelbaum *et al.* (2009) obtained a similar result in an approximate Fisher matrix prediction. Bridle & King (2007) showed that for future surveys the equation of state parameter w may be biased by order unity if IAs are significant and ignored.

Heymans *et al.* (2004*a*) used COMBO-17 measurements of the II signal from shear-shear correlations between galaxies close in redshift jointly with cosmic shear RCS and VIRMOS-DESCART shear-shear correlation data to constrain a simple intrinsic alignment model amplitude simultaneously with cosmological parameters. They marginalised over possible amplitudes of the II signal and found the fitted value of σ_8 was reduced by 0.03 relative to the value found when intrinsic alignments were ignored. Fu *et al.* (2008) estimated the amplitude of the GI signal marginalised over a range of σ_8, Ω_m values allowed by the cosmic shear data. They found it to be consistent with zero, and this conclusion held when various different scale ranges were used for each of the fits for cosmology or for the IA amplitude parameter. They concluded that they found no evidence for a non-zero GI signal. Schrabback *et al.* (2009) took significant steps to reduce the contamination of their cosmic shear constraints by IAs by excluding the auto-correlations for the 5 narrowest redshift bins and excluding luminous red galaxies (LRGs) from their analysis. They also showed results using only the autocorrelations, which were similar to those excluding the autocorrelations results, leading them to conclude that intrinsic alignments are not a significant contaminant.

Many current and future surveys are planned with cosmic shear as a major design driver,

in particular from the ground the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS)[1], the KIlo-Degree Survey (KIDS), the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS)[2] surveys, the Dark Energy Survey (DES)[3], the Large Synoptic Survey Telescope (LSST)[4], and space missions Euclid[5] and the Joint Dark Energy Mission (JDEM)[6].

The majority of observational constraints on galaxy IAs have been carried out at low redshift (Brown *et al.* 2002; Heymans *et al.* 2004*b*; Mandelbaum *et al.* 2006*c*) but these have recently been extended towards the redshifts relevant to cosmic shear surveys (Hirata *et al.* 2007; Mandelbaum *et al.* 2009). These studies have used spectroscopic galaxy surveys to calculate the galaxy position - galaxy shear correlation function and/or the galaxy shear - galaxy shear correlation function. Constraints calculated from these correlation functions have been fitted with simple models for IAs using a fixed cosmological model.

In this chapter we calculate constraints on cosmological parameters from cosmic shear data, taking into account the likely contamination from IAs for a range of IA models. We simultaneously fit an IA model and cosmological parameters to IA correlation functions. We then perform a joint analysis of cosmic shear and IA data, varying both cosmology and parameters within the IA model.

The chapter is organised as follows: Section 2.2 reviews the cosmic shear formalism and some basic models of IAs; Section 2.3 presents constraints from the shear-shear correlation data; Section 2.4 introduces an IA parameterisation and constraints from shear-position data; Section 2.5 uses shear-shear and shear-position data to produce joint constraints on IA and cosmological parameters and we conclude in Section 4.6.

Throughout we assume a flat, ΛCDM cosmology. When we refer to fiducial values for cosmological parameters we set the present day amplitude of linear mass fluctuations σ_8= 0.751, the baryon density $\Omega_{\rm b}$= 0.05, the matter density $\Omega_{\rm m}$= 0.3, and the Hubble parameter defined by $H_0 = 100h$ km s^{-1} Mpc^{-1} as h= 0.7 to match WMAP3. We assume a Harrison-Zel'dovich primordial power spectrum slope $n_s = 1$.

[1] http://www.cfht.hawaii.edu/Science/CFHLS/
[2] http://pan-starrs.ifa.hawaii.edu
[3] http://www.darkenergysurvey.org
[4] http://www.lsst.org
[5] http://sci.esa.int/euclid
[6] http://jdem.gsfc.nasa.gov

2.2 Cosmic Shear and intrinsic alignments

In this section we review the cosmic shear formalism before describing the origin of IAs and their effect on the measured cosmic shear signal. We then summarise three basic approaches to modelling IAs, beginning with the linear alignment model, then its non-linear extension and finally the halo model of Schneider & Bridle (2010*a*).

2.2.1 Cosmic Shear

Cosmic shear probes the expansion history and the growth of structure in the Universe using weak gravitational lensing. The effect is rather small ($\sim 1\%$) for most galaxies especially when compared to the intrinsic ellipticity of a single galaxy ($\sim 20\%$). Therefore it is necessary to average over many galaxies to detect a coherent signal. A convenient statistic is the Fourier transform of the two-point shear correlation function (ξ_E) between pairs of galaxies, the cosmic shear power spectrum C_l^{GG}. It is well approximated by projecting the three-dimensional matter power spectrum $P_\delta(k;\chi)$ at a redshift corresponding to a line-of-sight comoving distance χ, weighted by the lensing kernel, $W(\chi)$

$$C_l^{GG} = \int_0^{\chi_H} \mathrm{d}\chi \frac{W(\chi)W(\chi)}{\chi^2} P_\delta\left(k = \frac{l}{\chi}; \chi\right), \tag{2.1}$$

already introduced in equation 1.8.5, where in a flat universe

$$W(\chi) = \frac{3}{2c^2}\Omega_\mathrm{m} H_0^2 a^{-1}(\chi)\chi \int_\chi^{\chi_H} \mathrm{d}\chi_\mathrm{s}\, \mathrm{n}(\chi_\mathrm{s}) \frac{\chi_\mathrm{s} - \chi}{\chi_\mathrm{s}} \tag{2.2}$$

and a is the scale factor, χ_H is the comoving distance to the Hubble horizon, and the function $n(\chi_s)$ is the selection function of source galaxies per unit comoving distance, normalized as $\int d\chi\, n(\chi) = 1$.

The two-point shear correlation function is determined by the cosmic shear power spectrum via

$$\xi_E(\theta) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, C_l J_0(l\theta), \tag{2.3}$$

where J_0 is the zeroth order Bessel function of the first kind.

2.2.2 Intrinsic Alignments

If galaxies were randomly oriented on the sky then the cosmic shear signal would be relatively easy to observe because the random intrinsic contributions to the observed ellipticities would average

away to zero. However this is not expected to be the case. Galaxies form within large-scale gravitational potentials, producing some "intrinsic alignment" between galaxy shear. There are two ways in which IAs contribute to the observed shear power spectrum:

The II correlation: Intrinsic-Intrinsic galaxy alignments occur because physically close galaxies form in the same tidal field, causing an alignment of their halos or angular momentum vectors. As a result the galaxies point in the same direction. This alignment produces an increase in the measured shear power spectrum.

The GI correlation: Gravitational-Intrinsic alignments are a cross term between intrinsic ellipticity and cosmic shear. The intrinsic ellipticity of a galaxy is aligned with the density field in which it forms and this field in turn contributes to the lensing distortion of more distant galaxies. This double role causes an anti-correlation between galaxy ellipticities, because the closer galaxy may be expected to point towards the local overdensity while the distant galaxy is stretched tangentially around the overdensity. This leads to a suppression of the measured power spectrum.

The measured shear power spectrum C_l therefore arises from genuine cosmic shear, combined with IA terms produced by the II and GI correlations

$$C_l = C_l^{GG} + C_l^{II} + C_l^{GI}. \tag{2.4}$$

The strength of the II term depends strongly on the depth of the survey. A deep survey may contain galaxies at many different distances from the observer, so two galaxies which appear close on the sky are most likely to be far apart along the line of sight and therefore not physically close. Since the opposite is true for a shallower survey then the IA is stronger. In addition the lensing strength is smaller when all distances are reduced, so the cosmic shear contribution is small for a shallow survey.

The II and GI lensing power spectra are calculated analogously to C_l^{GG}

$$C_l^{II} = \int_0^{\chi_H} \frac{n^2(\chi)}{\chi^2} P_{\tilde{\gamma}^I}^{EE}(k,\chi)\mathrm{d}\chi \tag{2.5}$$

$$C_l^{GI} = \int_0^{\chi_H} \frac{2W(\chi)n(\chi)}{\chi^2} P_{\delta,\tilde{\gamma}^I}(k,\chi)\mathrm{d}\chi \tag{2.6}$$

where $k = l/\chi$, $P_{\tilde{\gamma}^I}^{EE}(k,\chi)$ and $P_{\delta,\tilde{\gamma}^I}(k,\chi)$ are the intrinsic-intrinsic and gravitational-intrinsic power spectra respectively, see Schneider & Bridle (2010*b*). The II term, Eqn. 2.5, depends on the galaxy redshift distribution, $n(\chi)$, while the GI term, Eqn. 2.6, depends on both $n(\chi)$ and the lensing weight function $W(\chi)$, as befits a cross-correlation.

In this chapter we illustrate the effect of IAs on cosmological constraints using data from the 100 Square Degree Weak Lensing Survey (Benjamin *et al.* 2007). We compare constraints from four different approaches to IAs:

1. Ignoring IA effects
2. The linear alignment model of IAs
3. The non-linear alignment (NLA) model of IAs
4. The Halo Model of IAs.

IAs can be studied with probes other than the usual cosmic shear measurement of shear-shear correlation functions. Section 2.4 below we use data from the Sloan Digital Sky Survey (SDSS) galaxy redshift survey shear-position correlation functions (Hirata *et al.* 2007) to constrain the amplitude, scale and luminosity dependence of IAs. These results can be combined with models of IA in cosmic shear predictions to produce less biased constraints on cosmological parameters. Details of the IA parameterisations are given below.

2.2.2.1 Linear Alignment Model

A popular simple method for modelling IAs is known as the Linear Alignment (LA) model because it assumes that the intrinsic ellipticity of galaxies is proportional to the curvature of the primordial large scale potential (Catelan *et al.* 2001; Hirata & Seljak 2004). This is thought to be most relevant for elliptical galaxies.

In the linear alignment model the II and GI power spectra are found to first order to be

$$P^{EE}_{\bar{\gamma}^I}(k,z) = \frac{C_1^2\bar{\rho}^2}{\bar{D}^2}P^{lin}_{\delta}(k,z) \tag{2.7}$$

$$P_{\delta,\bar{\gamma}^I}(k,z) = -\frac{C_1\bar{\rho}}{\bar{D}}P^{lin}_{\delta}(k,z) \tag{2.8}$$

where $P^{lin}_{\delta}(k)$ is the linear theory matter power spectrum, C_1 is a normalization constant, $\bar{D}(z) = (1+z)D(z)$ where $D(z)$ is the growth factor normalized to unity at the present day, and $\bar{\rho}(z)$ is the mean matter density of the Universe as a function of redshift (Hirata & Seljak 2004). This is a simple model which has been found to be relatively successful in the past (Mandelbaum *et al.* 2006*b*; Hirata *et al.* 2004). We take it that galaxies form in dark matter halos. These form via gravitational collapse and exhibit a triaxial profile. Whether halos are generally oblate or prolate

is an outstanding question. The model assumes that the major axis of the halo determines the ellipticity and is linearly related to the large-scale gravitational tidal field and that the galaxy ellipticities follow that of the DM halo. This model applies to elliptical galaxies. Spiral galaxies are thought to acquire ellipticity by tidal torquing. This is related quadratically to the tidal field but predicts no gravitational-intrinsic correlation.

2.2.2.2 Non-Linear Alignment Model

This linear alignment model is likely to hold on large scales but it takes no account of non-linear growth of structure. One somewhat ad-hoc method that has been employed (Bridle & King 2007) to rectify this deficiency is to substitute the non-linear matter power spectrum $P_\delta^{nl}(k)$, instead of the linear theory matter power spectrum $P_\delta^{lin}(k)$, into Eq. 2.7 and Eq. 2.8. In this work we use the Smith *et al.* (2003) prescription for the non-linear matter power spectrum. We call this approach the non-linear alignment (NLA) model.

2.2.2.3 Halo Model

The halo model of galaxy clustering (Scherrer & Bertschinger 1991; Scoccimarro *et al.* 2001) is a simple but effective predictor for galaxy clustering statistics. To first order in this model the Universe consists of dark matter halos, distributed according to linear theory. Each halo is described by a mass, drawn from a mass function, and a density profile. The forms of these functions are generally taken from averages of n-body simulations (see e.g. Cooray & Sheth 2002, for a review).

Applying the halo model to IAs provides a physically motivated way to model small-scale features of the IA signal. A simple example was provided in Schneider & Bridle (2010*a*) which we describe and use here. Each dark matter halo is taken to have a single central galaxy, positioned exactly at the centre of the spherical halo mass distribution. These central galaxies are assumed to have ellipticities determined by the large-scale density perturbations according to the linear alignment model. The satellite galaxy number density follows the density profile of the halo and their ellipticities are defined according to a distribution in the angle between the satellite major axis and the three-dimensional radius vector of the parent halo. Shear two-point correlation functions then consist of two components: a "one-halo term" arising from pairs of galaxies within the same halo, and a "two-halo term" due to pairs in different halos.

For IAs the one-halo term comes from the radial alignment of the satellite galaxies and the two-halo term is dominated by the correlation of the central galaxies in each halo. Within this

model one- and two-halo terms appear in both the II and GI IA power spectra

$$P^{EE}_{\tilde{\gamma}^I} = P^{EE,1h}_{\tilde{\gamma}^I} + P^{EE,2h}_{\tilde{\gamma}^I} \tag{2.9}$$

$$P_{\delta,\tilde{\gamma}^I} = P^{1h}_{\delta,\tilde{\gamma}^I} + P^{2h}_{\delta,\tilde{\gamma}^I}. \tag{2.10}$$

In the case of the II power spectrum, the two-halo term can be further broken up into terms due to the correlation of satellites with satellites, satellites with centrals and centrals with centrals, and, in the GI case, the two-halo power spectrum can be separated into a central and a satellite term. However, it was found (Schneider & Bridle 2010*a*) that the two halo terms are dominated by the central with central (II) and central (GI) terms, allowing us to ignore two-halo terms involving satellites. This greatly simplifies the equations, allowing us to write down the two-halo terms

$$P^{EE,2h}_{\tilde{\gamma}^I}(k,z) = \frac{C_1^2\bar{\rho}^2}{\bar{D}^2} P^{lin}_{\delta}(k,z) \tag{2.11}$$

$$P^{2h}_{\delta,\tilde{\gamma}^I}(k,z) = -\frac{C_1\bar{\rho}}{\bar{D}} P^{lin}_{\delta}(k,z) \tag{2.12}$$

i.e. the II and GI two-halo terms are just the same as the full linear alignment model of IAs.

Schneider & Bridle (2010*a*) provide fitting formulae for the one-halo power spectra

$$P^{EE,1h}_{\tilde{\gamma}^I}(k,z) = -\tilde{\gamma}^2_{\rm scale} \frac{(k/p_1)^4}{1+(k/p_2)^{p_3}} \tag{2.13}$$

$$P^{1h}_{\delta,\tilde{\gamma}^I}(k,z) = -\tilde{\gamma}_{\rm scale} \frac{(k/p_1)^2}{1+(k/p_2)^{p_3}} \tag{2.14}$$

where $\tilde{\gamma}_{scale}$ is the degree of alignment and determines the amplitude of the 1h power spectra, that could be set by comparison with simulations, and for the functions $p_i(z)$ (for i = 1,2,3) we use the fitting formulae provided in Schneider & Bridle (2010*a*), that are based on a halo model with redshift dependence determined by the Sheth-Tormen (Sheth & Tormen 1999) model for the halo mass function and bias as well as the linear growth function in their fiducial cosmology.

The IA power spectra $P^{EE}_{\tilde{\gamma}^I}$ and $P_{\delta,\tilde{\gamma}^I}$ enter both the shear-shear and shear-position calculations. In principle they depend on cosmology through the contents of the one- and two-halo terms, given in equations 2.11 to 2.14. The two-halo term contains the linear theory matter power spectrum and the growth rate, both of which depend on the values of cosmological parameters. The

one-halo term is derived from a consideration of the number halos as a function of halo mass, and a match to the redshift distribution of the survey in question. It was shown in Schneider & Bridle (2010*a*) that the dependence on the survey redshift distribution is weak and thus we neglect it here. The dependence of both terms on cosmology could be included, and this would make sense if we had complete faith in the models. We have the alternative of considering the IA power spectra for our fiducial model as a template representing a reasonable but poorly understood guess at the contribution from IAs. For this work we use the one-halo term as given in Schneider & Bridle (2010*a*) and ignore any cosmology dependence within it. We choose to fix σ_8 within the two-halo calculation to the fiducial value. Therefore the relative amplitude of the one- and two-halo terms stay roughly constant and the overall amplitude of the sum is varied (see below). This allows us to see IA impact, uncontaminated by varying σ_8 within part of the IA model. In this chapter we choose to vary the matter density inside the linear theory matter power spectra used in the two-halo terms in equations 2.11 and 2.12 and in the $\Omega_{\rm m}$ contribution to the normalised growth in the denominator of these equations.

To include the colour and luminosity dependence we generalise our halo model equations 2.9 and 2.10

$$P^{EE}_{\tilde{\gamma}^I} = \left(P^{EE,1h}_{\tilde{\gamma}^I} + P^{EE,2h}_{\tilde{\gamma}^I}\right)\left[A\left(\frac{L}{L_0}\right)^{\beta} f_r\right]^2 \tag{2.15}$$

$$P_{\delta,\tilde{\gamma}^I} = \left(P^{1h}_{\delta,\tilde{\gamma}^I} + P^{2h}_{\delta,\tilde{\gamma}^I}\right)\left[A\left(\frac{L}{L_0}\right)^{\beta} f_r\right] \tag{2.16}$$

which scales the amplitude of both one- and two-halo terms by the same factor, given in square brackets. The factor in square brackets is squared in the first equation and not in the second equation, to mimic a simultaneous modulation of the one-halo scaling parameter $\tilde{\gamma}_{\rm scale}$ and the two-halo amplitude parameter C_1. Throughout we retain the fiducial values of $\tilde{\gamma}_{\rm scale} = 0.21$ following Schneider & Bridle (2010*a*) and $C_1 = 5 \times 10^{-14}(h^2 M_\odot {\rm Mpc}^{-3})^{-1}$ following Bridle & King (2007).

L/L_0 is the normalised luminosity of each luminosity data bin as defined in Hirata *et al.* (2007), and f_r is the fraction of red galaxies in that data bin. This power law in luminosity is equivalent to that used in the power law fits of Hirata *et al.* (2007). The motivation for multiplication by the red fraction f_r is that we assume only red galaxies have IAs. These equations could be generalised to have different IA amplitudes for red and blue galaxies by adding a term proportional to the blue fraction $(1 - f_r)$ with a different variable amplitude parameter. However, we defer such modelling to future work. A and β are free parameters. Note that the above equations

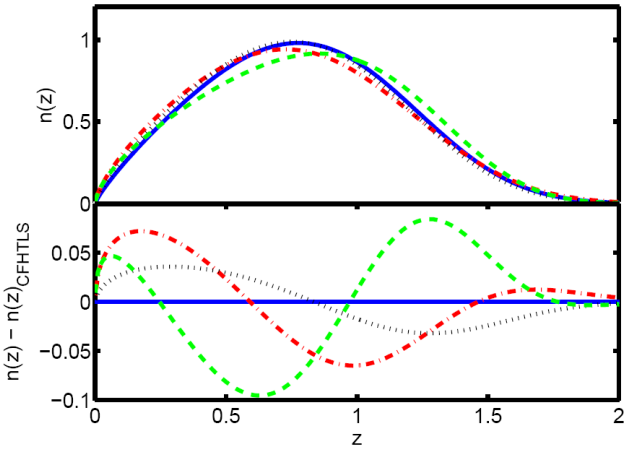

Figure 2.1: The upper panel shows galaxy redshift distribution, n(z), for the four surveys comprising the 100 deg^2 cosmic shear dataset. These are best fits to the Smail-type redshift distribution equation taken from Benjamin *et al.* (2007), hence the smooth curves. The lower panel shows the $n(z)$ difference with respect to CFHLTS, $n(z) - n(z)_{\rm CFHTLS}$. CFHTLS is the solid line, GaBoDS is dot-dashed, RCS is dotted and VIRMOS-DESCART is dashed.

reduce to the basic halo model for $A = f_r = 1$, $\beta = 0$.

Throughout section 2.3 we assume that the IA model is perfectly known. This certainly overestimates our knowledge of this effect and can lead to the over-estimation of cosmological parameters as IAs are treated as extra signals rather than noise. We begin to rectify this is section 2.4 by marginalising over a number of parameters in the IA model.

2.3 Shear-shear correlations

In this Section we summarise the cosmic shear data we use and compare it with predicted correlation functions with the various IA models. We then make a first calculation of the impact of the models on constraints on the amplitude of matter clustering and the matter density of the Universe.

2.3.1 Cosmic Shear Data

The dataset used to constrain models in this chapter is the 100 Square Degree weak lensing survey (Benjamin *et al.* 2007). This combines data from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS)-Wide, the Garching-Bonn Deep Survey (GaBoDS), the Red-sequence Cluster Survey (RCS) and the VIRMOS-DESCART surveys.

The CFHTLS-Wide data included in this compilation covers an area of 22 deg^2, reaching a depth of i' = 24.5 (Hoekstra *et al.* 2006). There is 13 deg^2 of data from the GaBoDS survey

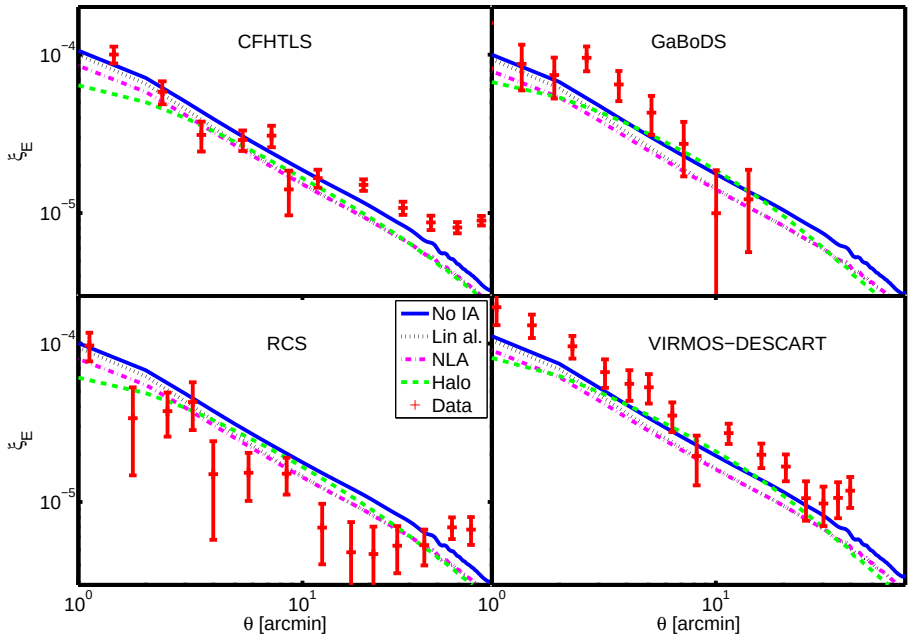

Figure 2.2: Shear correlation function, ξ_E, data from the four surveys comprising the 100 deg^2 cosmic shear dataset along with correlation functions for each of the four IA models calculated using the same fiducial cosmology for all but the appropriate n(z) for each survey. The correlation function without IAs is shown by the solid line, the linear alignment model as dotted line, the NLA model as dot-dashed and the halo model as the dashed line. The data, with 68% errors, appears as crosses.

which uses objects which lie in the interval R $\in$ [21.5,24.5] (Hetterscheidt *et al.* 2007). The RCS data covers 53 deg^2 with a limiting magnitude of 25.2 in the R_C band (Hoekstra *et al.* 2002). The VIRMOS-DESCART data has an effective area of 8.5 deg^2 and a limiting magnitude of $I_{AB} = 24.5$ (Le Fèvre *et al.* 2004). Errors and covariances are taken from Benjamin *et al.* (2007).

Throughout this chapter we use the redshift distributions for each data set as given by Eq. 9 of Benjamin *et al.* (2007) using the parameters given in the upper section of their Table 2 for the high confidence regime (fitted to photometric redshifts in the range $0.2 < z < 1.5$ from Ilbert *et al.* 2006). Benjamin *et al.* (2007) took into account uncertainties in the photometric redshift distributions by sampling from tuples of these parameter values and selecting the best fit values for each of the three parameters. For computational practicality we here ignore uncertainties in the photometric redshift distributions. We therefore next describe in more detail the relative differences between the redshift distributions of the different surveys, ignoring the uncertainties.

Since the impact of IAs on each dataset hinges so much on the redshift distribution we show

the redshift distributions and the difference in distribution, with respect to CFHTLS, for each dataset in Fig. 2.1. Note that the n(z) profiles are based on fits made by Benjamin *et al.* (2007) to a Smail-type (Smail & Dickinson 1995) n(z) profile. This profile is defined by three parameters: α, β and z_0. For each of the four constituent surveys this parameter space was Monte Carlo sampled 1000 times. We have taken the best-fit parameter combination for each survey, hence the smooth curves in Fig. 2.1 describe analytic functions which approximate the true n(z) distribution of each survey. We see that VIRMOS-DESCART and GaBoDS have the greatest number of low redshift galaxies and VIRMOS-DESCART shows a signifigant excess at $z \sim 1.4$.

2.3.2 Correlation Functions Including Intrinsic Alignments

Fig. 2.2 shows the correlation functions of the four constituent surveys of the 100 deg^2 dataset, along with 67% errors, as red crosses. This information is taken from the appendices of Benjamin *et al.* (2007). We overlay the predicted correlation functions for each of the IA models we consider. It may appear that the differences in correlation function between surveys is much greater than the difference produced by using the various IA models. However, one should note that the plotted correlation functions are calculated for our fiducial cosmology, with no attempt to find the best fit to each survey. In the following analysis each survey is analysed separately, using the best-fit cosmology suitable for that survey. Once this is done the difference between IA models becomes statistically significant. Note that Fig. 1 in Benjamin *et al.* (2007) plots similar calculated correlation functions but uses a different, best-fit, σ_8 for each survey.

Overall the GI contribution dominates the IA signal and has the effect of reducing the shear-shear correlation signal; hence all the correlation functions which include an IA component are mostly lower than the equivalent correlation function without IAs (solid line). As expected, the halo model (dashed line) agrees with the LA (dotted line) and NLA (dot-dashed line) models at large angular separation where the two-halo term and linear theory dominate. The slight difference present between the halo and LA models in GaBoDS and VIRMOS-DESCART disappears at larger angular separation than shown in the figure. The halo and LA models diverge at small angular separation when the additional one-halo term becomes important and the shear-shear correlation amplitude is reduced due to the negative contributions from the GI term. The NLA model is intermediate between the LA and halo model at small angular separations. The increase in the halo model correlation function at intermediate scales is caused by an increased II contribution on these scales. In the halo model of IAs the galaxies align towards the center of dark matter halos, and as a result are also aligned with each other. Therefore the one-halo term for the GI

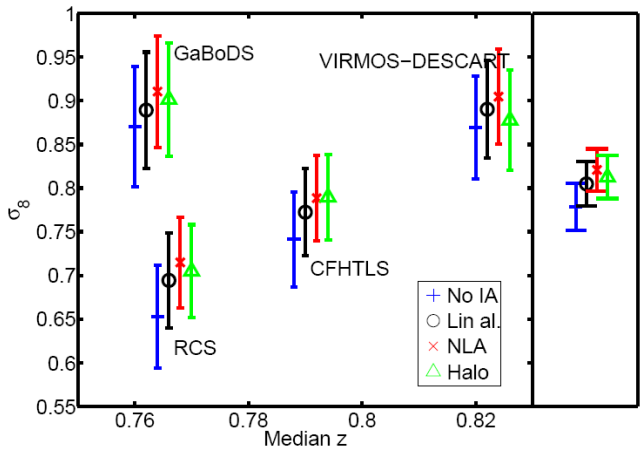

Figure 2.3: 68% confidence limits on the amplitude of fluctuations σ_8 for each survey within the 100 deg^2 cosmic shear dataset. From left: GaBoDS, RCS, CFHTLS, VIRMOS-DESCART, and combined. Each IA model is shown separately for each survey, plotted at the median redshift of each survey. From left: No IAs (horizontal crosses), linear alignment (circles), NLA model (diagonal crosses) and halo model (triangles). IA points are offset from the median redshift for clarity. The combined constraints for the whole 100 deg^2 dataset are shown in the panel on the right. This analysis varies σ_8 leaving other parameters fixed at their fiducial values, including $\Omega_{\rm m} = 0.3$.

signal dominates on a scale corresponding to the typical distance of galaxies from the halo center, whereas the one-halo term for the II signal occurs at the typical halo diameter. This explains why the GI term dominates at small scales and the II term has a bigger impact at intermediate scales.

We now compare the predicted correlation functions for the different surveys. The only difference between the predictions comes from the different redshift distributions used. The IA signals dominate over the cosmic shear signal at low redshift. A large number of physically close galaxies produces a large II effect. Surveys with a large number of low redshift galaxies will have a large proportion of these physically close strongly aligned galaxies. Fig. 2.1 shows that VIRMOS-DESCART (dashed line) and GaBoDs (dot-dashed line) have the largest proportion of low redshift galaxies, so we expect the strong increase at intermediate scales in Fig. 2.2. Conversely CFHTLS has the lowest proportion of low-z galaxies and we see little increase in its correlation function due to an II contribution.

2.3.3 Impact on σ_8 constraints

We are interested in the biasing of cosmological constraints when IAs are erroneously ignored. We compare predicted and observed correlation functions and produce χ^2 values at a range of

amplitudes of matter clustering, parameterised by σ_8. These are used to calculate a best fit σ_8 and corresponding 68% confidence limits. This was done for each of the four cosmic shear datasets separately and the analysis was repeated using the four different IA approaches. The results are shown in Fig. 2.3. The right-hand panel shows the 68% confidence limits using the combined 100 deg^2 data, obtained by multiplying together the probabilities $\Pr(\sigma_8) \propto \exp(-\chi^2/2)$ from each survey. In this analysis, all other cosmological parameters are held at their fiducial values and we fix $A = 1$, $\beta = 0$ and $f_R = 1$. In all analysis quoted in this chapter, we kept σ_8 constant, at its fiducial value, within the IA models themselves which can be scaled by an overall amplitude parameter A. As usual, the cosmic shear correlation function has a strong dependence on σ_8.

The results show a dependence of the measured σ_8 on the IA model. Ignoring IAs (horizontal/vertical crosses) produces results which are consistently below the results when IAs are taken into account, irrespective of the IA model used. This behaviour is consistent with the correlation functions seen in Fig. 2.2. The dominance of the GI term has the effect of lowering the overall shear-shear correlation function prediction meaning that constraints using an IA model favour a larger value of σ_8 to compensate for the reduced amplitude. The effect is more pronounced for the NLA model (diagonal crosses) than for the linear alignment model (circles) because suppression of the correlation function due to the GI term is stronger in the NLA. When the halo model is used (triangles) the best fit σ_8 is larger than that when IAs are ignored, but the difference can be smaller than for the NLA or LA models due to the contribution from the II correlations at intermediate separations which partially cancel the GI terms. We note that the difference between the halo model constraint and that when IAs are ignored is least for GaBoDS and VIRMOS-DESCART. This is consistent with their relatively large number of low-redshift galaxies. This produces an increased II contribution at intermediate scales, lowering constrained σ_8 values. The effect, the reduction in difference from No IA when the halo model is used, is biggest for VIRMOS-DESCART because it has the greatest constraining power at these intermediate scales.

Although IAs are an important cosmic shear systematic and a clear biasing is apparent, it is encouraging to note that the effect is not catastrophic: the best-fit values for almost all surveys and IA models lie within the 68% confidence limits of the results in which IAs are ignored. Because the uncertainties on σ_8 are smaller for the joint constraints (right hand panel of Fig. 2.3), the total biases are now similar to the 68 per cent confidence limits. We use a more sophisticated model for how IAs affect cosmic shear constraints in Section 2.5.

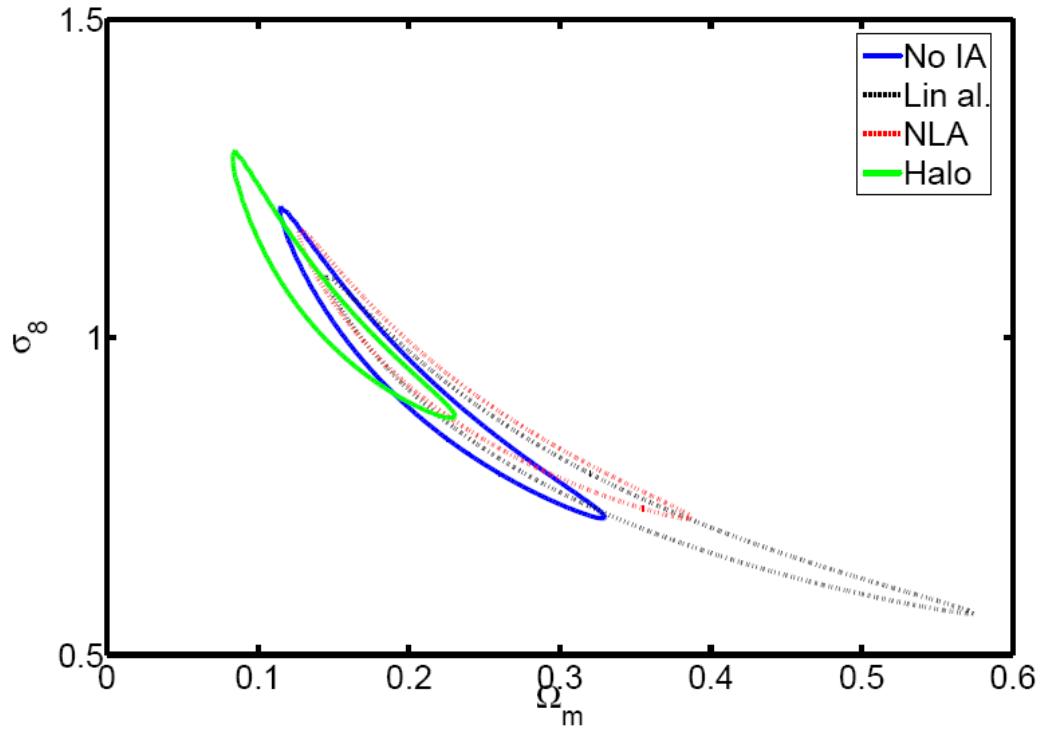

Figure 2.4: 68% Confidence limits in σ_8 - $\Omega_{\rm m}$ parameter space combining the four surveys in the 100 deg^2 cosmic shear dataset for each IA model. Constraints where IAs are ignored in the modelling are shown by the dark, solid (blue) contour; constraints assuming the linear alignment model are shown by the dark dotted (black) contour; the NLA model by light dotted (red) and the halo model by the light solid (green) contour. IA parameters are set to $A = 1$, $\beta = 0$, $f_r = 1$. All other cosmological parameters were fixed at their fiducial values.

2.3.4 Impact on constraints on σ_8 and $\Omega_{\rm m}$

We next calculate constraints from the cosmic shear data allowing both the fluctuation amplitude σ_8 and the matter density $\Omega_{\rm m}$ to vary. Other cosmological parameters are held at their fiducial values. As previously noted, σ_8 is fixed within the IA models but $\Omega_{\rm m}$ is allowed to vary throughout, in the lensing calculation and all IA models. The results are shown in Fig. 2.4 as 68% confidence contours in σ_8-$\Omega_{\rm m}$ parameter space. The likelihood surface is calculated on a grid of points in σ_8-$\Omega_{\rm m}$ with a spacing of 0.03 in σ_8 and 0.06 in $\Omega_{\rm m}$. Any greater sampling density was prohibitevly expensive in terms of computational time. Axis limits of Fig. 2.4 represent the edges of the grid. All results for the various IA models show the usual cosmic shear degeneracy between σ_8 and $\Omega_{\rm m}$. As expected, at a fixed $\Omega_{\rm m} = 0.3$ we can read off the σ_8 constraints shown in the previous figure. However the relative pattern between models depends on the assumed value of $\Omega_{\rm m}$ and furthermore different $\Omega_{\rm m}$ ranges are preferred by different models.

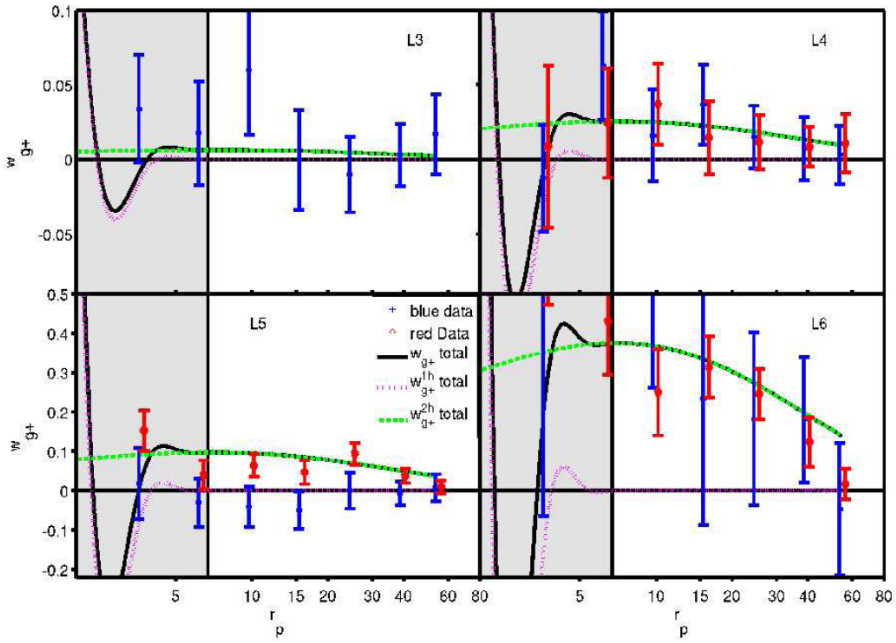

Figure 2.5: Comparison of the shear-position correlation function, w_{g+}, from the Hirata *et al.* (2007) data and the halo model prediction using fiducial cosmology, A = 1, $\beta = 1.44$, $L_0 = L_4$. Predicted correlation functions are shown for the one-halo term only (dotted), two-halo term only (dashed) and the total (solid). Blue galaxy data for each luminosity bin is plotted as crosses, red galaxies as diamonds. The shaded region shows distance scales excluded from our analysis.

The most interesting result is that from the halo model of IAs (light solid (green) contour), which favours the high σ_8 and low $\Omega_{\rm m}$ end of the degeneracy. Very roughly speaking the shear-shear correlation function data provide two pieces of information: an amplitude and a slope. The amplitude constrains a degenerate combination of σ_8 and $\Omega_{\rm m}$ but the slope of the correlation function partly breaks this degeneracy and determines the preferred range range of $\Omega_{\rm m}$ along the degeneracy line. The most constraining part of the data is CFHTLS at large scales where the data points are far from the predictions (this has by far the largest contribution to the χ^2). The effect of the halo model is to steepen the correlation function on large scales. The steeper halo model correlation function at large scales is particularly favoured by the data in the low $\Omega_{\rm m}$/high σ_8 region. Increasing $\Omega_{\rm m}$ steepens the correlation function, this is disfavoured by the data when using the halo function as shown by the new position of the likelihood contour.

2.4 Shear-position correlations

In this Section we introduce a new data set which was produced with the aim of measuring the IA signal. We use the galaxy shear - galaxy position two-point function that quantifies whether the shape of one galaxy points at the position of another galaxy. If we have a good understanding of how the galaxy positions are related to the mass distribution then this effectively constrains the IA power spectrum $P_{\delta,\tilde{\gamma}^I}$. We compute predictions for the two point function using the halo model for IAs, then compare it to the data to measure IA model parameters.

2.4.1 Data

We use data measured from the SDSS galaxy survey, a 5-band (ugriz) photometric survey over π steradians of the sky, with spectroscopic follow-up of 10^6 objects (York *et al.* 2000). We use the galaxy position - galaxy shear correlation functions from the data described in Mandelbaum *et al.* (2006*c*) and calculated in Hirata *et al.* (2007). These are calculated from the main spectroscopic sample (z = 0.05 - 0.2).

This is split by luminosity and colour and each of these subsamples is used to trace the intrinsic shear field. The luminosity subsamples span L3 (one magnitude fainter than L_*) to L6 (two magnitudes brighter). Each of these is divided into "red" and "blue" colour subsamples, defined using an empirically determined, redshift dependent separator, $u - r = 2.1 + 4.2z$ in the observer frame. The full sample is used to trace the density field. Ellipticity measurements were taken from Mandelbaum *et al.* (2005).

2.4.2 Correlation Functions

The correlation function is calculated by taking pairs of galaxies where one galaxy is from the shear catalogue (luminosity and colour bin) and the other is from the density catalogue (which is a composite of all the luminosity and colour bins). The ellipticity of the galaxy from the shear catalogue is calculated in a coordinate system aligned with the vector between the two galaxies in the pair. If the galaxy in the shear catalogue is aligned along the vector joining the pair then this number will be positive. If it points perpendicular to this line then the ellipticity in this coordinate system will be negative. These ellipticities are averaged over all pairs, as a function of separation on the sky r_p and the separation along the line of sight Π, taking into account edge effects using a generalized Landy-Szalay (Landy & Szalay 1993) estimator. This is integrated over the range $-60 < \Pi/h^{-1}\mathrm{Mpc} < 60$. We use the covariance matrices between bins in angular separation

which were provided.

We predict the shear-position correlation function w_{g+} using

$$w_{\delta+} = -\frac{1}{2\pi} \int P_{\delta,\tilde{\gamma}^I}(k) J_2(kr_p)\, k\, \mathrm{dk} \tag{2.17}$$

(Hirata *et al.* 2007, Eq. 23) which we assume is related to the galaxy correlation function by the galaxy bias b_g

$$w_{g+} = b_g w_{\delta+} \tag{2.18}$$

where we use galaxy bias estimates from Table 3 of Hirata *et al.* (2007) and we allow for the variation of galaxy bias with σ_8 as they describe. Essentially this correlation asks does the ellipticity of one galaxy point towards another galaxy in a pair? As galaxies are (biased) tracers of the underlying denisty distribution, this can be thought of as probing the shear-density correlation at a given separation. We are interested in the impact of IA models on cosmological constraints. We neglect errors on galaxy bias in this analysis for simplicity. The error on the galaxy bias is about 10% (see Hirata et al. 2007) and this would correspondingly impact on the width of the σ_8 constraints from the shear-position correlation functions. As expected, and as can be seen from Fig. 7, this would not have a significant impact on our analysis because the main constraint in the σ_8 direction comes from the shear-shear information. We use the generalised halo model as given in Eq. 2.15 and Eq. 2.16 and take values for the luminosity of each SDSS shear-position correlation function bin from the minimum magnitude in each bin as given in Hirata *et al.* (2007) Table 1. We set L_0 to be the luminosity of the L4 bin. We set the red fraction f_r to be unity for the red samples and zero for the blue samples, so that effectively we do not use the blue samples in the constraints. Following Hirata *et al.* (2007), we use only the data points for separations greater than $4.7h^{-1}$ Mpc where the bias can reasonably be assumed to be linear. The overall amplitude scaling parameter A and luminosity power law β are the two free parameters we consider.

Fig. 2.5 shows the shear-position correlation function, w_{g+}, for each luminosity and colour bin in the Hirata *et al.* (2007) data (crosses for blue galaxies, diamonds for red galaxies). Note that this data and associated erroes has been taken from table A1 of Hirata *et al.* (2007). Correlations between error bars are accounted for in our likelihood analysis using the correlation matrices from table A2 of Hirata *et al.* (2007).

Also shown are predictions using the halo model, using for illustration $A = f_r = 1$ and $\beta = 1.44$, the best-fit β value provided by the third row of the lower half of table 4 in Hirata *et al.* (2007). Note that this is a more realistic value than the $\beta = 0$ used previously but it will act to

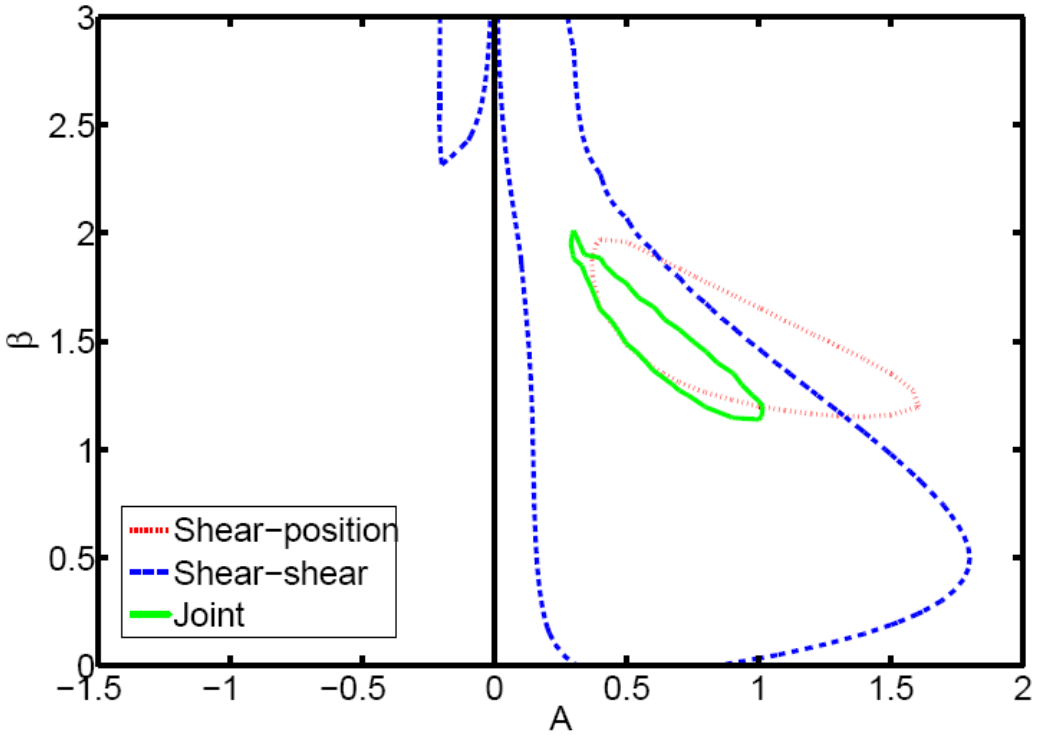

Figure 2.6: 68% confidence limits on A and β, the amplitude and luminosity dependence of the IA power spectra. Constraints are plotted from the Hirata *et al.* (2007) shear-position data (dotted contour), the 100 deg^2 shear-shear data (dashed contour) and joint constraints (solid contour). σ_8 was marginalised over and all other cosmological parameters were fixed at their fiducial values.

reduce the amplitude of the IA signal. Separate correlation functions are plotted for the one-halo contribution (dotted line), two-halo contribution (dashed line) and the total (solid line). As we would expect, the two-halo term provides the bulk of the GI signal at high separation, and falls off at small separation. An interesting feature of the one-halo correlation function is that as the separation is decreased it rises from zero, then falls to negative values, before rising sharply at low separation. This behaviour was discussed in Schneider & Bridle (2010*a*) with reference to their Fig. 2 in their Section 2A. The grey-shaded regions on Fig. 2.5 show the scales which have been excluded from our analysis of the IA data, following the practice of Hirata *et al.* (2007). We see that essentially only the two-halo term is significant in the fits to this dataset.

2.4.3 Constraints on Intrinsic Alignment Parameters

In this section we use the shear-position correlation function data to constrain A and β, the free parameters in the IA model. The dotted line in Fig. 2.6 shows the resulting 68% confidence

limits. Both shear-shear and shear-position data are marginalised over σ_8. We chose to pivot the luminosity power law at the SDSS L4 luminosity bin (L_0 corresponds to the magnitude of the L4 bin). The red galaxies are either in luminosity bin L4 or higher luminosity bins, therefore an increase in β will mostly increase the predicted IA signal relative to the data. The same is true for A, therefore we expect them to be anti-correlated, as seen in Fig. 2.6. We discuss the remaining lines on this figure in the following section, but note here that the shear-position correlation data is by far the most important dataset when constraining these two parameters alone.

Our results agree well with those of Hirata *et al.* (2007) who find $\beta = 1.44^{+0.63}_{-0.62}$ when using all the SDSS Main data (blue and red) simultaneously, with the same minimum separation in the correlation function. By construction the best fit A is around unity, since the vast majority of the constraint comes from the two-halo term of the halo model which corresponds to the linear theory prediction estimated from Hirata & Seljak (2004) which was fitted to the Heymans *et al.* (2004*a*) data which gave a similar amplitude as the Mandelbaum *et al.* (2006*c*) result.

To allow a comparison between the cosmic shear data and the IA model used for the galaxy survey data values of f_r and L were calculated as a function of redshift for each component survey of the 100 deg^2 dataset using a conditional luminosity function model based on Cooray (2006) and outlined in appendix 2.7.

In the previous section we computed cosmological constraints from cosmic shear data with and without consideration of IAs. The "No IA" results correspond to $A = 0$ on this Figure, that is clearly ruled out by the SDSS data. The "halo model" results corresponded to $A = 1$, $\beta = 0$, $f_r = 1$ which is also ruled out by this figure but would be appropriate if all the galaxies in the cosmic shear survey had the same luminosity and colour as the SDSS Red L4 sample. In the following Section we present a full joint analysis of both samples.

2.5 Joint shear-shear and shear-position constraints

In this Section we compare constraints on parameters from shear-shear and shear-position correlation functions and combine the constraints to produce limits on both IA and cosmological parameters. To make our previous analysis of the cosmic shear data more general it would be good practice to allow the IA free parameters A and β to vary, then marginalise over them rather than assuming they were fixed to the basic halo model prescription. This would inevitably lead to a loss of constraining power. However, this loss can be made less severe by calculating joint constraints on IA parameters from cosmic shear and galaxy survey data. With this approach, the extra

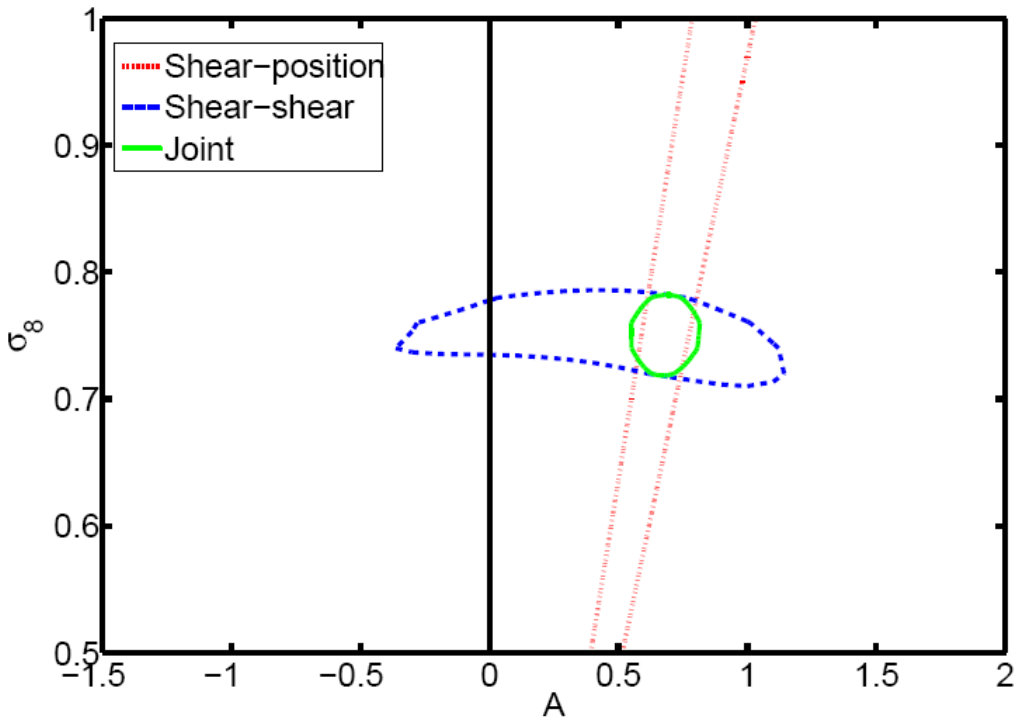

Figure 2.7: 68% confidence limits on σ_8 and A, the IA power spectrum amplitude parameter. Constraints are plotted from the Hirata *et al.* (2007) shear-position correlation function data (dotted lines), the Benjamin *et al.* (2007) 100 deg^2 shear-shear correlation data (dashed contour) and joint constraints (solid contour). The IA luminosity dependence power law slope was fixed $\beta = 1.44$, and all other cosmological parameters were fixed at their fiducial values.

information about IAs provided by galaxy surveys is used to produce less biased cosmological constraints from cosmic shear.

2.5.1 Joint Constraints on Intrinsic Alignment Parameters

We combine constraints from the two datasets by simply adding together their χ^2 values. This is acceptable given that the surveys do not overlap significantly. We calculate χ^2 values from both datasets as a function of both IA parameters and cosmological parameters. For the shear-shear dataset the IA parameters enter via the IA power spectra (Eq. 2.15 and Eq. 2.16) which are projected onto the sky (Eq. 2.5 and Eq. 2.6) and added to the cosmic shear contribution to produce the full shear-shear power spectrum (Eq. 2.4) and the correlation function is calculated by Eq. 2.3. The shear-position correlation functions are computed from the shear-density correlation function (Eq. 2.18) which is a inverse Hankel transform (Eq. 2.17) of the relevant IA power spectrum (Eq. 2.6). The bias values given in Hirata *et al.* (2007) are calculated from measurements of

galaxy clustering (position-position correlation functions) and therefore depend on the assumed value of σ_8. We take this into account to obtain a bias values for each σ_8 value considered. As discussed at the end of Section 2.2.2, we fix σ_8 in the IA power spectra and allow Ω_m to vary in the linear theory matter power spectrum contribution to the two-halo term.

First we allow only σ_8 and A, the amplitude of the IA signal, to vary, with the rest of the cosmological parameters set to their fiducial values and a fixed luminosity dependence power law slope $\beta = 1.44$ (the best-fit value given in Hirata *et al.* 2007). The results are shown in Fig. 2.7 as 68% confidence contours in σ_8-A parameter space. Contours are shown for the shear-position correlation functions (nearly vertical lines) shear-shear correlation functions (nearly horizontally elongated contour) and the combined constraint (roughly at the intersection). The shear-position correlation function constraints on the IA amplitude parameter A for the fiducial σ_8 value are as expected from Fig. 2.6 for the fixed β value, with A significantly larger than zero. In the analysis of the shear-position correlation function data σ_8 is held fixed in in the calculation of all our IA models. However, the variation of galaxy bias as a function of σ_8 produces the expected degeneracy with A. Increasing σ_8 above the fiducial value decreases galaxy bias, requiring a greater A to compensate and vice-versa.

The shear-shear correlation functions do themselves place some constraint on the IA parameter A, preferring a range of order unity. A negative value of A would correspond to galaxies pointing in the opposite direction to that used in the standard models. The linear alignment model (two-halo term) would contain galaxies pointing perpendicular to the tidal stretching expected by the gravitational potential curvature. The one-halo picture would contain galaxies which are aligned tangentially to the center of the halo. We described earlier a rough picture of shear-shear correlation function constraints in which essentially the data measure the amplitude and slope of the correlation function. The amplitude essentially fixes a degenerate combination of σ_8 and A, and the constraint on A must therefore come from the shape of the correlation function. For the fiducial $\Omega_{\rm m}$, Fig. 2.7 tells us that a large IA contribution to the shear-shear power spectra distorts them too much. This can be understood by examining Fig. 2.2, in which the data points fit well with the shape of the lensing-only ("No IA") predictions, whereas the halo model predictions tend to be more curved over the scales probed.

The direction of degeneracy between A and σ_8 from shear-shear information alone can also be understood in terms of the shear-shear correlation function datapoints in Fig. 2.2. In general there is a balance of effects between the II and GI contributions. If we imagine a universe in which the GI term did not exist but the II term did, an increase in the IA amplitude parameter A would

add power to the predicted shear-shear correlation function. To keep the predictions consistent with the data points it would be necessary to decrease σ_8 to reduce the lensing contribution. This would give a negative correlation between A and σ_8, seen for larger positive A values. In this unphysical universe containing only II terms, the direction of degeneracy would appear reversed for negative A because A appears as a squared quantity in the II terms. Indeed we do see the direction reversed in this figure. In the physical Universe the effect of the GI contribution is to break the symmetry slightly, and prefer slightly more positive A values since it causes a slight cancellation in the IA effect for positive A values which fits better to the shape of the correlation functions that an addition of power.

The great complementarity of the two datasets is most clearly illustrated in this figure. The main constraints on cosmology come from the shear-shear correlation data and the main constraints on IAs from the shear-position data. The joint constraints focus at the intersection between the two relatively degenerate constraints, as expected.

The analysis was repeated, allowing A, β and σ_8 to vary. We marginalised over σ_8 to produce the results shown in Fig. 2.6. All other parameters were fixed at their fiducial values. The shear-shear correlation function data (dashed) provides much weaker constraints in the A-β plane- partly due to marginalisaion over σ_8 in the power spectrum but primarily because the cosmic shear data is not broken into luminosity bins, so has very little power to constrain β. Even so we can distinguish the same degeneracy direction as present in the shear-position data and positive A is favoured at 68% confidence. Most of the constraining power in the joint constraints (solid) comes from the shear-position data (dotted).

2.5.2 Joint Constraints on Cosmology

Here we revisit the constraints on cosmological parameters from Section 2.3.2 showing how they change when we allow the IA parameters (A and β) to vary and be marginalised over. Joint shear-shear and shear-position constraints are also presented to show constraints on cosmology when IAs are taken into account as well as possible given current models and data.

Fig. 2.8 shows constraints on the amplitude of matter fluctuations σ_8 for each shear-shear survey separately as a function of survey median redshift, and also the joint constraint for all shear-shear surveys in the right hand panel. The blue crosses in Fig. 2.8 are the same as those in Fig. 2.3 and show constraints on σ_8 when IAs are ignored. In Fig. 2.3 we illustrated the effect of IAs on the different IA models using a simple version of the IA power spectra in which $A = 1$, $\beta = 0$ and we assumed $f_r = 1$. In Fig. 2.8 we use only the halo model and use the values for L

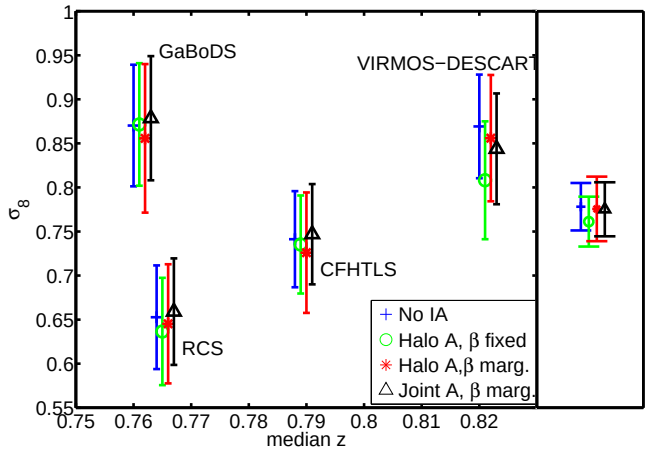

Figure 2.8: 68% confidence limits on σ_8 for each survey within the 100 deg^2 shear-shear correlation function dataset (from left: GaBoDS, RCS, CFHTLS, VIRMOS-DESCART and combined). For each survey constraints shown, from left: without IAs (crosses); using the halo model with $A = 1$, $\beta = 1.44$ (circles); halo model, marginalised over A and β (stars); joint shear-shear and shear-position constraints, marginalised over A and β (diamonds). The right-side panel shows joint constraints from the four surveys. All other cosmological parameters were fixed at their fiducial values.

and f_r as a function of redshift as discussed in Appendix 2.7. For the circles in Fig. 2.8 we fixed $A = 1$, $\beta = 1.44$, and we see that the σ_8 results using the halo model are all now lower than before. The result of using the more realistic L and f_r values instead of $\beta = 0$, $f_r = 1$ is to lower all the IA contributions. This is to be expected since f_r is always less than or equal to unity in the more realistic model, and in addition f_r is particularly small at low redshifts where faint satellite galaxies dominate and the IA power spectra are largest. The biggest modification to the predicted correlation functions is the reduced GI contribution, especially on smaller scales. Therefore the total shear-shear correlation is increased and the fitted σ_8 is reduced to compensate. The trends with different shear-shear surveys reflect those already discussed in reference to Fig. 2.3. As a result, the joint constraint (right panel of Fig. 2.8) shows the halo model with $A = 1$, $\beta = 1.44$ (circle) is now slightly lower than when IAs are ignored. This is consistent with the contours shown in Fig. 2.7 which have a lower σ_8 at $A = 1$ than $A = 0$.

We next allow A and β to vary within the halo model, and we marginalise over both parameters with flat wide priors using only the shear-shear correlation data. The results are shown by the stars in Fig. 2.8. As expected, the uncertainties are larger than when these parameters are kept fixed. The best fit values are about the same as when no IAs are included. We would expect this from Fig. 2.7 in which the $A = 0$ line cuts a cross-section through the banana-shaped degeneracy contour from

shear-shear correlations alone which is roughly in the middle of the σ_8 range allowed by the whole contour.

Joint constraints from the combined cosmic shear and IA data (triangles) gives tighter constraints on σ_8 than from cosmic shear data alone, as expected. The constraint from all shear-shear surveys (right hand panel) is consistent with the contours in Fig. 2.7, where it is clear that the joint constraint (solid line) gives a very similar σ_8 value to the cosmic shear constraint without IAs (a slice through the cosmic shear contour at A=0).

The variation between different IA treatments are most strongly pronounced for the case of the VIRMOS-DESCART survey (highest median redshift of the four survey points in Fig. 2.8). Although it has the highest median redshift, we have seen from Fig. 2.1 that it also contains a large fraction of low-redshift galaxies and hence most physically close galaxies. This increases the II IA signal, that is significant for the range of scales probed by VIRMOS-DESCART. This boosts the amplitude of the predicted shear-shear correlation function points, requiring lower σ_8 to compensate. This behaviour produces the larger bias observed in VIRMOS-DESCART on going from no IA to the halo model.

The new σ_8 value can be higher or lower than when IAs are ignored (horizontal crosses), depending on the dataset in question. However all the new cosmology constraints (triangles) lie easily within the 68% confidence limits of the constraint without IA, including the constraint for all shear-shear surveys combined.

We repeated the analysis for constraints in σ_8-$\Omega_{\rm m}$ space. All other cosmological parameters were held at their fiducial values. Fig. 2.9 shows the 68% confidence contours for cosmic shear data without IAs (blue solid contour), that are the same as the dark solid lines in Fig. 2.4. The halo model with fixed IA parameters (green dotted contour) shows similar behaviour to that discussed in Fig. 2.4, favouring the high σ_8 and low $\Omega_{\rm m}$ region. Marginalising over the IA amplitude and luminosity-dependence parameters, A and β (red dotted contour), relaxes the constraint, extending the degeneracy further into the high σ_8, low $\Omega_{\rm m}$ region. The joint constraints (black solid contour) using both shear-shear and shear-position information show a preference for low σ_8 and high $\Omega_{\rm m}$ but are still consistent with our fiducial cosmology. The preference for higher $\Omega_{\rm m}$ values occurs because, as noted in Hirata *et al.* (2007), the shear-position data favours slightly steeper correlation functions than predicted by a standard fiducial cosmology (see also the L6 panel of Fig. 2.5). Increasing $\Omega_{\rm m}$ produces a correlation function that falls off faster at larger scales because the linear theory matter power spectrum peak is moved to the right (smaller scales) on increasing $\Omega_{\rm m}$ due to the earlier epoch of matter-radiation equality.

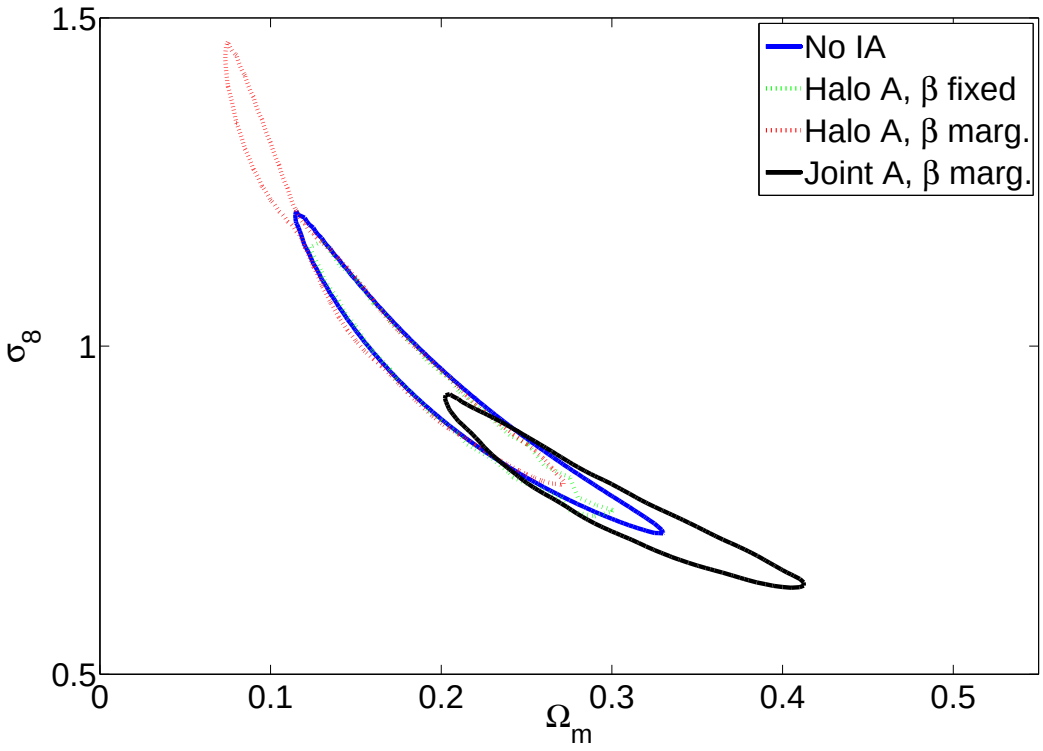

Figure 2.9: Combined 68% confidence limits in σ_8-$\Omega_{\rm m}$ parameter space for the 100 deg^2 shear-shear correlation function dataset. Constraints shown without IAs (blue solid contour); using the halo model with $A = 1$, $\beta = 1.44$ (green dotted contour); halo model, marginalised over A and β (red dotted contour); joint shear-shear and shear-position constraints, marginalised over A and β (black solid contour). All other cosmological parameters were fixed at their fiducial values.

This new constraint on cosmology from cosmic shear self-consistently takes into account IA effects and we find $\Omega_{\rm m} = 0.18 \pm 0.09$, $\sigma_8 = 0.88 \pm 0.18$ at 68 % confidence, as opposed to the results in which IAs are ignored, that give $\Omega_{\rm m} = 0.20 \pm 0.10$, $\sigma_8 = 0.86 \pm 0.21$. When we constrain simultaneously with shear-shear and shear position data we find $\Omega_{\rm m} = 0.29 \pm 0.09$, $\sigma_8 = 0.74 \pm 0.13$, after marginalising over IA parameters. There is a strong degeneracy between σ_8 and $\Omega_{\rm m}$ when using current cosmic shear data alone, so it is better to look at the full two-dimensional contours (Fig. 2.9) or at cross-sections. When we fix $\Omega_{\rm m} = 0.3$ our best self-consistent shear-shear result including IAs (right hand panel of Fig. 2.8) is $\sigma_8 = 0.75 \pm 0.04$ as compared to that when IAs are ignored of $\sigma_8 = 0.75 \pm 0.03$. A joint constraint from shear-shear and shear-position data, marginalising over IA parameters, gives $\sigma_8 = 0.75 \pm 0.03$. The "pinch" feature in Fig. 2.9 is an artefact of the sampling resolution of our grid in $\sigma_8 - \Omega_m$ space. The particular position of smaples near that portion of the field can cause the error ellipse to be drawn tighetr/looser than a more smoothly sampled field would produce. Our ability to sample at higher resolution was

limited by constraints of computing power.

2.6 Conclusions

IAs are expected to be an important source of systematic error in cosmic shear measurements if completely ignored. Several previous constraints on cosmology from shear-shear correlation functions have considered the possible levels of contamination from IAs, but the recent constraints presented have not included the impact of IAs (Benjamin *et al.* 2007; Fu *et al.* 2008; Schrabback *et al.* 2009). These surveys are therefore referred to as "cosmic shear" surveys when, infact, they measure cosmic shear contaminated by IAs.

In this chapter we use three physically motivated IA models, of increasing complexity, to demonstrate how constraints on the amplitude of fluctuations σ_8 and matter density $\Omega_{\rm m}$ are changed depending on the assumptions about IAs. The size of the change in σ_8 depends on the redshift distribution of galaxies in the shear-shear survey and on the range of scales probed. In particular this affects the interplay between the two different types of IA (II and GI) which have competing effects on cosmological constraints. A simple examination of IAs which assumes all galaxies are like the SDSS Main Red L4 sample suggests that σ_8 has previously been underestimated by about one standard deviation. The underestimate appears to be larger using the non-linear matter power spectrum in the linear alignment model for IAs than when using the linear matter power spectrum. On using the halo model for IAs the bias depends significantly on the range of scales probed by the survey due to the larger effect of II at intermediate scales which biases the measured σ_8 downwards when taken into account.

We have performed the first constraints on multiple IA model parameters from observed shear-position correlation functions using a physically motivated model for IAs. We also show the first simultaneous constraints on IA and cosmological parameters from either shear-position or shear-shear correlation functions. As expected, the majority of the constraint on cosmology comes from the shear-shear correlation functions, and the majority of the constraint on IAs comes from the shear-position correlation functions.

We have used a motivated model for luminosity and colour evolution of shear-shear galaxy samples as a function of redshift to improve our constraints on cosmology from shear-shear correlation data in the presence of the halo model for IAs. In general this reduces the overall effect of IAs on cosmology, which we attribute to the decreased contribution from IA at low redshifts. We consider constraints on the amplitude of fluctutions σ_8 from shear-shear data alone with

fixed IA amplitude and luminosity dependence parameters, and compare this with constraints after marginalising over these parameters. As expected the constraints are weaker after marginalisation, but are not biased significantly. This is due to the relatively flat degeneracy between σ_8 and the IA amplitude arising from the competing effects of GI and II terms, and the ability of shear-shear information alone to place some constraint on the IA amplitude if other cosmological parameters are held fixed.

The model we used for the luminosity and colour evolution of the shear-shear galaxy sample could be improved in its complexity, for example by allowing a population of faint blue galaxies at high redshift. Furthermore we assumed the shear-shear galaxy sample had a single luminosity and colour at each redshift, whereas a more sophisticated analysis would integrate over ranges in each.

A joint analysis of shear-shear and shear-position correlation functions has been discussed as a promising tool for removing IAs from future cosmic shear datasets (Zhang 2008; Bernstein 2009; Joachimi & Bridle 2009). We perform the first joint analysis of these correlation functions on observational data. The uncertainties on the amplitude of fluctuations σ_8 are now reduced to the size found when IAs are ignored. The estimated value of σ_8 is either decreased or increased relative to the case where IAs are ignored. However, for the combined constraints from the Benjamin *et al.* (2007) shear-shear data compilation the overall σ_8 value is barely changed on performing the full joint analysis of shear-shear and shear-position correlation functions with IAs, relative to the conventional analysis in which IAs are ignored and only shear-shear information is used.

The constraints in the two-dimensional parameter space of σ_8 and $\Omega_{\rm m}$ show a bigger change, that we have attributed to the slope of the shear-position correlation function preferring higher $\Omega_{\rm m}$ values. This raises the question of how to incorporate cosmological parameters into IA predictions. For example we could have considered the IA power spectra from our fiducial cosmological model as template functions which are allowed to vary only in amplitude. It also places a new burden of accuracy on measurements of the shear-position correlation function, that in turn depend on measurements of galaxy biasing for their interpretation.

We have not repeated the full analysis of Benjamin *et al.* (2007), who consider several other potential sources of systematic uncertainty including photometric redshift and shear calibration errors. We defer this to future investigations that could incorporate the very latest cosmic shear data. The recent re-analysis of the COSMOS cosmic shear data (Massey *et al.* 2007) by Schrabback *et al.* (2009) investigates a number of potential cosmic shear systematics including IAs. The photometric redshifts are used to produce the first cosmic shear tomographic cross-correlation

functions which allow a much better control of IAs. Furthermore they investigate the effect of removing red galaxies from the cosmic shear sample which should help to reduce the effect of IAs further. This would have to be carefully modelled in the predictions of the IA contribution to the shear-shear correlation functions.

We believe that future cosmic shear analyses should include a model for IAs and use the existing galaxy survey data to help constrain this simultaneously with cosmology, as set out in this chapter. We have taken separate cosmic shear data (from Benjamin *et al.* (2007)) and shear-position data (from Hirata *et al.* (2007)) using their stated correlation functions and errors. In the future we hope to work from the shear maps and position catalogues of shear surveys to produce constraints that explicitly include IAs and the full galaxy position and position-shear correlation information. In chapters 3 & 4 we present forecasts for future cosmic shear surveys based on this approach.

The modelling in this chapter could be improved by extending the flexibility in the IA model, for example by allowing the one- and two-halo terms to vary independently in amplitude, and allowing an unknown amplitude for the IAs for blue galaxies. Furthermore the amplitude parameters could be free functions of redshift and even scale. The IA models we have used specify simultaneously the two types of IA, II and GI. However, a more general model of IAs can lead to independent uncertainties in each of the source terms of these effects (Bernstein 2009, see the bias and correlation terms of the intrinsic convergence with the lensing potential in).

However, if such a large degree of flexibility were allowed with current shear-position information then the constraining power of current cosmic shear data would be significantly reduced. Therefore it is important to obtain better measurements of the shear-position correlation functions, and/or improve models of IAs to reduce the number of free parameters. In this analysis we have used measurements of the galaxy shear-position correlation function made from a spectroscopic galaxy redshift survey. It should be possible to make these measurements from the same data used for the cosmic shear survey itself, however this has not yet been done. This will require a careful understanding of the survey mask as well as a measurement of the galaxy biasing to allow the shear-position correlation function to be related to the shear-density correlation function. In addition there are constraints in the literature from the shear-shear correlation functions of low redshift galaxy samples as well as for higher redshift spectroscopic galaxy samples which can be used to provide further constraints on the IA contributions.

2.7 Appendix- Models for mean luminosity and red galaxy fraction

This appendix describes the models we use to describe the luminosity distribution and fraction of red/blue galaxies in each of the surveys we consider. It is believed that the importance of IAs will depend heavily on this galaxy properties and it will be essential to include this sort of information in future analysis of cosmic shear data. We apply these models to our cosmic shear and galaxy position information in section 2.5.

To match the luminosity and colour dependence of the IA model in the cosmic shear data we use a model for the conditional luminosity function (CLF) closely following Cooray (2006). The conditional luminosity function $\Phi(L|M,z)$ gives the mean number of galaxies in a dark matter halo of mass M at redshift z with luminosity in the range $[L, L+dL]$. This function provides a way to add luminosity dependence to the halo model description of galaxy statistics.

Following the usual practice in the halo model, the CLF model is constructed from separate models for the central and satellite galaxies. The first step is to specify the mean luminosity of the central galaxy in a halo of mass M, $L_c(M,z)$, which we model as a broken power-law in mass according to Eq. 3 of Cooray (2006); where the mean luminosity of the central galaxy increases quickly with mass for low halo masses and then flattens off for larger halo masses. The total luminosity of a halo should include contributions from satellite galaxies for masses greater than some threshold, $M_{\rm sat}$,

$$L_{\rm tot}(M,z) = \begin{cases} L_c(M,z) & M \leq M_{\rm sat}(z) \\ L_c(M,z)\,\left(\frac{M}{M_{\rm sat}(z)}\right)^{\beta_{\rm s}(z)} & M > M_{\rm sat}(z) \end{cases} \tag{2.19}$$

where $\beta_s = 0.8$, $M_{\rm sat}(z) \equiv M_{\rm sat}\,{\rm erf}(5\,z)$ and $M_{\rm sat} = 10^{13}\,h^{-1}M_\odot$. This particular z dependence is an ad-hoc adjustment to allow lower mass halos to host satellites at low z.

In the simplest model the CLF for central galaxies is just a delta-function centered at L_c, although Cooray (2006) suggests a "rounded" step to allow for scatter in the minimum mass halo hosting a galaxy,

$$\Phi^{\rm cen}(L|M,z) = \Theta_H\,(M - M_{\rm cen-cut})\ \text{Log-Normal}(L|L_c,\sigma_c), \tag{2.20}$$

where $\sigma_c = 1.15$, Θ_H is the Heaviside step function and $M_{\rm cen-cut} = 10^{10}\,h^{-1}M_\odot$.

The satellite CLF is modeled as a power-law in luminosity with maximum luminosity equal to half the central galaxy luminosity (so that satellites do not dominate the total luminosity of the halo),

$$\Phi^{\rm sat}(L|M,z) = \Theta_H\left(\frac{L_c}{2} - L\right)\,A(M,z)\,\frac{1}{L}. \tag{2.21}$$

The amplitude, $A(M,z)$, is determined by requiring $\int_{L_{\rm min}}^{L_{\rm max}} dLL\Phi^{\rm sat}(L|M,z) = L_{\rm tot} - L_c$, giving,

$$A(M,z) \approx \frac{L_{\rm tot}(M,z) - L_c(M,z)}{\min\left(L_{\rm max}, L_c(M,z)/2\right) - L_{\rm min}(z)} \tag{2.22}$$

For modeling the luminosity dependence in the 100 sq. deg. lensing data sets, we set the integration limits $L_{\rm min}, L_{\rm max}$ to match the apparent magnitude limits given in table 1 of Benjamin *et al.* (2007). Note the luminosity integration limits become redshift dependent when matching fixed apparent magnitude limits. For the upper integration limit we further set $L_{\rm max}(M,z) = \min\left(L_c(M,z)/2, L_{\rm max,survey}(z)\right)$.

The HOD functions giving the number of central and satellite galaxies in a halo of mass M are obtained by integrating the CLFs over luminosity,

$$N_g^{\rm cen,sat}(M,z) = \int_{L_{\rm min}}^{L_{\rm max}} dL\, \Phi^{\rm cen,sat}(L|M,z). \tag{2.23}$$

Integrating over halo mass, we can relate the CLFs to the redshift distribution of the galaxy sample as,

$$\begin{aligned}\frac{dN}{dzd\Omega}(z) = \chi^2(z)\,\frac{d\chi}{dz}(z)\,\int dM\, n(M,z) \\ \times\left(N_g^{\rm cen}(M,z) + N_g^{\rm sat}(M,z)\right),\end{aligned} \tag{2.24}$$

where $\chi(z)$ is the comoving distance as a function of redshift[7]. We match the redshift distributions from Benjamin *et al.* (2007) with the CLF prediction to fit the (redshift-dependent) normalization of the central galaxy luminosity $L_c(M,z)$.

With the parameters of our CLF model matched to the magnitude limits and redshift distribution of each lensing survey, we calculate the mean luminosity of each galaxy sample as,

$$\langle L\rangle(z) = \frac{\int dM\, n(M,z) \int dL\, L\, \Phi(L|M,z)}{\int dM\, n(M,z) \int dL\, \Phi(L|M,z)}. \tag{2.25}$$

We show our model for $\langle L\rangle(z)$ for each of the four lensing surveys in Fig. 2.10.

We again follow Cooray (2006) to model the fraction of red central galaxies as a function of halo mass as a "rounded" step function,

$$f_{\rm red,cen}(M,z) = f_{\rm red,cen}(z)\frac{1}{2}\left(1 + {\rm erf}\left(\frac{\ln(M) - \ln(M_{\rm cen})}{\sigma_{\rm red,cen}}\right)\right). \tag{2.26}$$

This approaches the constant value $f_{\rm red,cen}(z)$ as M becomes large and goes to zero for small M. Note this has no luminosity dependence.

[7] assuming zero spatial curvature.

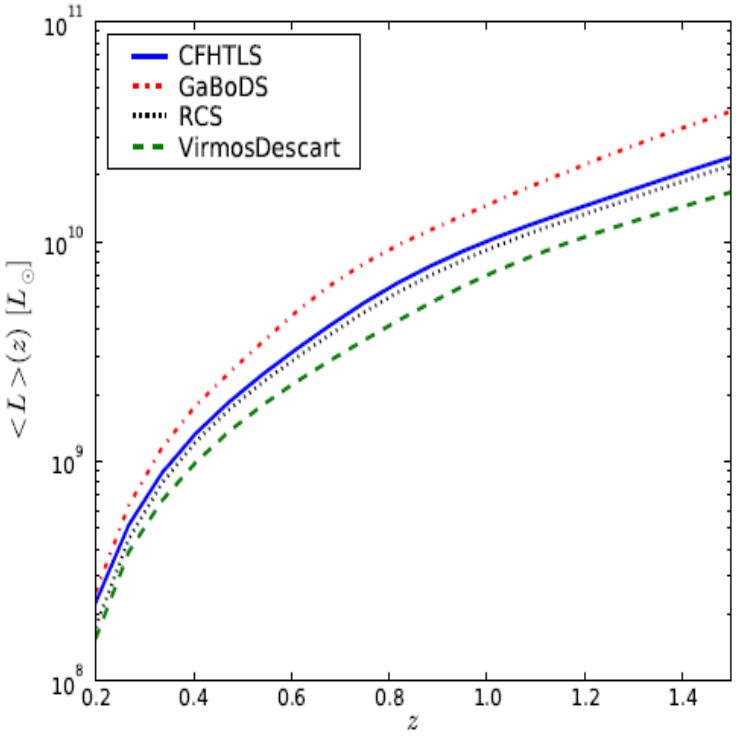

Figure 2.10: Mean luminosity model versus redshift for each of the 4 lensing surveys.

The fraction of red satellites in a halo of mass M is modeled as,

$$f_{\rm red,sat}(M, L, z) = g_{\rm red,sat}(z)\,(g(M, z) + h(L, z)) + f_{\rm red,sat}(z) \tag{2.27}$$

where $g(M, z)$ and $h(L, z)$ are "rounded step functions" in mass and luminosity, respectively. For low masses and luminosities, the red satellite fraction is the constant $f_{\rm red,sat}(z)$ while for high masses and luminosities it is the different constant $2g_{\rm red,sat}(z) + f_{\rm red,sat}(z)$. For simplicity, we set $f_{\rm red,cen}$, $f_{\rm red,sat}$, and $g_{\rm red,sat}$ to the z-independent constants 0.6, 0.4, and 0.2.

From these pieces, the fraction of all red galaxies as a function of redshift is,

$$\begin{aligned} f_{\rm red}(z) = \frac{1}{n_(\rm gal)} \int dM\, n(M, z) \int dL \\ \times [\Phi^{\rm cen}(L|M, z) f_{\rm red,cen}(M, z) \\ + \Phi^{\rm sat}(L|M, z) f_{\rm red,sat}(M, L, z)], \end{aligned} \tag{2.28}$$

where $n_(\rm gal)$ is the comoving number density of galaxies in the sample.

Our models for $f_{\rm red}(z)$ for each of the 4 lensing surveys are shown in fig. 2.11. Ignoring redshift evolution in $f_{\rm red,cen}$ and $f_{\rm red,sat}$ is almost certainly incorrect and, in combination with the

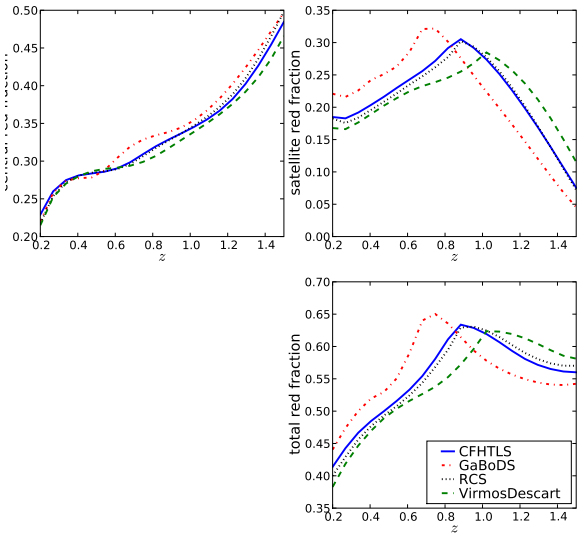

Figure 2.11: Red galaxy fraction versus redshift for each of the 4 lensing surveys. The top panels show the red fraction for central and satellites separately in our halo model while the bottom panel shows the red fraction for all galaxies.

fixed magnitude cutoffs in our model, tends to select excess red galaxies at high redshift (when normalized to give reasonable red fractions at low redshift). Because the IA halo model includes contributions only from red galaxies, we consider this model for $f_{\rm red}$ to give an upper bound on the possible IA contamination to the weak lensing data. This is suitable for the goal in this chapter of assessing the impact of IAs as systematic errors in lensing surveys. To improve the predictive power of our IA halo model in the future it will be important to better understand the redshift evolution of the red fraction of galaxies.

CHAPTER 3

THE EFFECT OF INTRINSIC ALIGNMENTS ON CONSTRAINING MODIFIED GRAVITY WITH COSMIC SHEAR

3.1 Introduction

We have discussed how gravitational lensing arises from the distortion of galaxy images due to gravity as the light traverses space between emission and observation. Upcoming imaging surveys promise to provide wide area measurements of both the distribution of luminous matter, through galaxy position surveys, and the distribution of the total matter distribution, through weak lensing shear measurement. Ground-based surveys include Kilo-degree Survey (KIDS), Panoramic Survey Telescope & Rapid Response System (Pan-STARRS) project, Subaru Meaurement of Images and Redshifts (SuMIRe), Dark Energy Survey (DES) and the Large Synoptic Survey Telescope (LSST). There are two proposed space missions: Euclid and Wide-Field Infrared Survey Telescope (WFIRST). The expectation is that these will support a significant improvement in the constraints on dark energy.

Our standard cosmological picture is often called ΛCDM cosmology. The 'CDM' stands for cold dark matter which is invisible but gravitates like normal matter. The 'Λ' is the dark energy component, otherwise known as a cosmological constant because it takes the same form as the cosmological constant term introduced by Einstein to maintain a static universe. The inclusion of an unusual negative pressure fluid, making up $\sim$75% of the energy budget of the Universe is, in some ways, an admission of ignorance by cosmologists. To rectify the situation much effort

has been focused on constraining some properties of the dark energy equation of state (Albrecht *et al.* 2006; Peacock *et al.* 2006) in the hope that we may be able to produce a satisfactory model which explains the origins and behaviour of dark energy. One major priority is to test whether dark energy is a pure cosmological constant or whether its equation of state varies with time. Future cosmic shear surveys have been designed to help reach goals on dynamic dark energy constraints.

There has been an active discussion of the implications of such measurements for constraining cosmological models that might predict growth histories that are distinct from those of the'vanilla' ΛCDM cosmological model, in particular those arising from large scale modifications of gravity (MG). Such modifications of gravity have been proposed as alternative ways to address the problem of explaining the accelerating expansion of the universe. Rather than considering a specific theory, one can assess the ability of surveys to constrain modifications to gravity using a phenomenological parameterization to allow deviations of how the gravitational metric perturbations are related to the underlying matter distribution and motion. We can write the metric in the form for the conformal Newtonian gauge:

$$\mathrm{d}s^2 = -(1+2\Psi)\mathrm{d}t^2 + a^2(1-2\Phi)\mathrm{d}\mathbf{x^2}, \tag{3.1}$$

where Ψ is the perturbation of the time-part of the metric Φ is the perturbation of the space-part. Such a parameteriztion is then well suited to be tested by weak lensing measurements, as lensing is sensitive to the sum of the two potentials, while galaxy position is sensitive to the Newtonian potential, Ψ, only.

We have shown in chpater 2 how using cross-correlation between galaxy position and lensing observables can greatly increase constraining power and play an important role in mitigating critical sources of systematic error, in particular the effect of galaxy intrinsic alignments on the observed convergence signal (Joachimi & Bridle 2010; Kirk *et al.* 2010; Zhang 2010). It has been suggested that the cross-correlations may be particularly important for constraining modified gravity theories (Zhang *et al.* 2007; Zhang 2008).

The Intrinsic Alignment (IAs) systematic arises when galaxies are aligned in their orientations. However a model for alignments is not well developed, which means one cannot well characterize a contribution from this systematic error. The IA effect is the strongest astrophysical systematic contaminant of the cosmic shear signal. Implementing a realistic IA model and marginalising to remove the bias from cosmological parameters can result in a significant reduction in constraining power (Joachimi & Bridle 2009; Kirk *et al.* 2010; Bridle & King 2007; Joachimi *et al.* 2011; Joachimi & Schneider 2008). The degeneracy between IAs and modified gravity (MG) parame-

ters has not been explored before. In this work we investigate the potential impact that intrinsic alignments could have on our ability to constrain modifications to gravity.

When a parameterized model of IA is considered, it can be imagined that since both MG and IA alter the convergence signal there might be a degeneracy between MG and IA parameters. Here, we characterize and quantify this degeneracy, and assess how it deteriorates our ability to set constraints on modifications to General Relativity. Joachimi & Bridle (2009) implemented the suggestion of Zhang (2010) and demonstrated that the inclusion of galaxy position information went some way towards mitigating the impact of intrinsic alignments on dark energy constraints. We expand this treatment to constraints on modified gravity, defining our own MG Figure of Merit (FoM) for easy comparison of dark energy and MG constraining power.

As more data becomes available, we will require a more accurate characterisation of the systematic effects contaminating our cosmic shear signal. We present a reformulated, expanded version of the LA model (Hirata & Seljak 2004), building on work by Joachimi & Bridle (2009). The effects of MG and galaxy clustering at non-linear scales are not well understood. We include a cut in scale designed to minimise the impact that these uncertainties could have on our constraints.

In section 3.2 we outline a new formalism used to describe intrinsic aligments and include them in the lensing and galaxy position correlations including the effects of modified gravity. Section 3.3 describes our fiducial survey specifications and the Fisher Matrix formalism we use to predict constraints on cosmological parameters. We present our results in 3.4 and conclude in 3.5.

3.2 Formalism

In this work, we adopt a cosmology consistent with the best fit WMAP 7 year results (Komatsu *et al.* 2011) ($\Omega_b h^2$= 0.00223, $\Omega_c h^2$= 0.0111 Ω_K=0, A_s=2.20,n_s=0.969,$w_0 = 1$, $w_a = 0$). We treat the background evolution as fixed and determined by the standard cosmology, as any viable modified gravity model must reproduce observed expansion history, so that we do not expect to derive new constraints on MG parameters from expansion history. We now describe the observables used and how we compute them, the intrinsic alignment model and how it is incorporated into the weak lensing signal, and briefly describe our MG model.

3.2.1 Observables

We are interested in how well measuments of the weak lensing shear distortions, and galaxy positions, from future surveys will be able to constrain deviations from standard gravity as described

above. Each of these observables is key to detecting changes in the growth history of the universe as they depend differently on the metric potentials.

We characterize the observations by their 2D angular power spectrum, C_l, for auto and cross correlations between each observable. For two fields X and Y, under the Limber approximation, C_l is given by

$$C_l^{XY} = \int_0^{\chi_\infty} \frac{d\chi}{\chi^2} W_X(\chi) W_Y(\chi) T_X(k_l, \chi) T_Y(k_l, \chi) \Delta_R^2(k_l), \tag{3.2}$$

where X, Y=g and G for galaxy position and lensing shear fields respectively, W_X is the window function associated with the field X, and T_X is the transfer function defined below and $k_l = (l + 1/2)/\chi$.

The transfer and window functions for ISW, galaxy and weak lensing are well known. The transfer functions for galaxy and weak lensing sources are

$$T_g = \tilde{\Delta}_c, \tag{3.3}$$

$$T_G = -\frac{k^2}{2}(\tilde{\Phi} + \tilde{\Psi}), \tag{3.4}$$

where the transfer function T_X (or T_Y) is normalized so that $X^2(k, \chi) = T_X^2(k, \chi)\Delta_{\mathcal{R}}^2(k)$ where $\Delta_{\mathcal{R}}^2(k)$ is the dimensionless primordial spectrum of curvature fluctuations. These transfer functions encode the physics which modifies the primordial power spectrum, both in terms of scale dependent behaviour e.g. the turnover in the matter power spectrum, and growth of structure with redshift. Each type of observable will have its own transfer function, for example the galaxy position transfer function will have to include a bias term. This transfer function notation is useful because it relates everything back to the primordial spectrum.

The galaxy and lensing window functions are dependent on the distribution of galaxy number density in each redshift bin i for the relevant survey, $n^i(\chi)$, and

$$W_g^i(\chi) = n^i(\chi), \tag{3.5}$$

$$W_G^i(\chi) = \int_\chi^{\chi_\infty} d\chi' n^i(\chi') \frac{r(\chi) r(\chi' - \chi)}{r(\chi')}, \tag{3.6}$$

where n^i is normalized such that, $\int_0^{\chi_\infty} d\chi n^i(\chi) = 1$.

In relating the observed correlations of galaxies to the dark matter correlation functions above, we must account for a bias between dark and luminous matter. This bias is dependent on the galaxy type, redshift and environment in a way that is poorly understood. We allow for this uncertainty by introducing redshift and scale dependent bias parameter, b_g, to relate the autocorrelations, and an independent, correlation parameter r_g, to describe the bias in the cross correlation of luminous

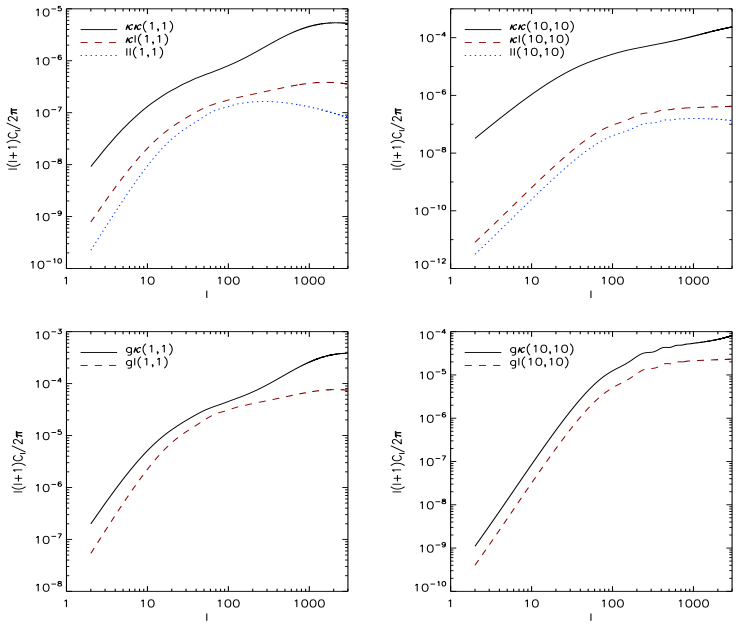

Figure 3.1: Angular power spectra showing different contributions from intrinsic alignments. Upper: Shear-shear power spectra. Solid lines show the total observable power spectra, dashed lines show the GI contribution and dotted lines show the II contribution. Lower: Shear-position power spectra. Solid lines again show the total observable power spectra, and this time the dashed lines show the gI intrinsic alignment contribution. Left: Correlation of the lowest redshift tomographic redshift bin with itself. Right: Correlation of the highest redshift tomographic bin with itself.

matter and the dark matter,

$$P_{gg} = b_g^2 P_{\delta\delta} \tag{3.7}$$

$$P_{g\delta} = b_g r_g P_{\delta\delta} \tag{3.8}$$

where $P_{\delta\delta}$ is the matter power spectrum (see e.g. Dekel & Lahav 2000 for a review, or see the equations set out in the context of cosmic shear in Bernstein 2009). Generally, b_g andr_g can vary with scale and redshift. We define our parameterisation in 3.2.4 below.

3.2.2 Intrinsic Alignments

Cosmic shear describes the distortion of the image of a distant galaxy due to the bending of light from that galaxy by gravity as it passes massive large-scale structure. We analyse our galaxy sample tomographically, cutting it into slices in redshift and treating the auto- and cross-correlation of these slices. For a galaxy in the i^{th} z-bin, the observed ellipticity, ϵ, of the galaxy can be written as a sum of three independent contributions: the cosmic shear γ_G, the intrinsic, non-lensed shape of the galaxy, γ_I, and apparent ellipticity introduced through instrumental and foreground noise, ϵ_{rnd},

$$\epsilon^{(i)}(\theta) \quad = \quad \gamma_G^{(i)}(\theta) + \gamma_I^{(i)}(\theta) + \epsilon_{rnd}^{(i)}(\theta). \tag{3.9}$$

The cosmic shear signal, γ_G is very small, around 0.03 and we cannot measure directly the intrinsic shear of any individual galaxy. To recover the cosmic shear, therefore, one can average over a number of galaxies on a small patch on a sky. Assuming that their intrinsic ellipticities are distributed randomly and that their light has passed by the same large scale structure so that they acquire the same shear signal, on averaging the intrinsic ellipticities cancel and we are left with the cosmic shear signal.

In reality the assumption that intrinsic ellipticities are randomly distributed on the sky is not accurate. They may be aligned by the process of galaxy formation, for example they may be tidally stretched to point at a nearby dark matter clump. This results in two different types of effect on the shear two point correlation function. The Intrinsic-Intrinsic (II) alignment applies to physically close galaxies. These form in the same large-scale gravitational potential and their intrinsic ellipticities tend to align with the curvature of that potential. When these galaxies are observed on the sky they will tend to point in the same direction. This alignment produces a spurious correlation which adds to the observed cosmic shear signal.

Somewhat more subtle is the Gravitational-intrinsic (GI) alignment which applies to galaxies close on the sky but separated in redshift. The intrinsic ellipticity of the foreground galaxy will tend to align with the nearby gravitational potential which, in turn, is responsible for gravitationally lensing the background galaxy. This tends to produce an anti-correlation which subtracts from the observed cosmic shear signal as we observe the galaxies oriented orthogonally.

We use the linear alignment (LA) model of Hirata & Seljak (2004), including the erratum of 2010. We therefore can write down the intrinsic alignment source term

$$I(k,z) = \frac{C_1 k^2}{4\pi G}\frac{a(z)}{D(z)}\Psi(k,z), \tag{3.10}$$

where $\Psi(k,z)$ is related to the matter overdensity via the Poisson equation

$$\Psi(k,z) = -\frac{4\pi G}{k^2}\rho(z)a^2(z)\delta(k,z). \tag{3.11}$$

This source term formalism lets us conveniently write the angular power spectrum as

$$C_l^{XY} = \int_0^{\chi\infty} \frac{d\chi}{\chi^2} W_X(\chi) W_Y(\chi) I_X(k,z) I_Y(k,z), \tag{3.12}$$

a slight modification of Eqn. 3.2 where we are able to take a more physical view of the mechanism which sources the IA (and other) terms.

However, while the above approach is consistent with Hirata & Seljak (2004) and the subsequent use of the Linear Alignment model in the literature, it neglects certain physical ideas implicit in the LA model approach. Principally, the idea that galaxy intrinsic alignments are frozen in at some early time (loosely called primordial in Hirata & Seljak (2004)) and do not evolve with redshift. In fact the original LA model was designed so that most redshift evolution (e.g. by the growth, $D(z)$) cancelled out but it did retain a factor of $\rho(z)a(z)$ as appears in Joachimi & Bridle (2009) eqn. 37. This evolution with redshift is seemingly contrary to the principles Hirata & Seljak (2004) claim drive the LA model and indeed, the correction published in 2010 produces equations in which the redshift dependence cancels out.

We recast the LA formalism in a more consistent approach, asserting that the IA source term, $I(k,z)$, is entirely determined by physics at some redshift of galaxy formation, z_f . We then have

$$I(k,z_f) = \frac{C'_1 k^2}{4\pi G}\frac{a(z_f)}{D(z_f)}\Psi(k,z_f), \tag{3.13}$$

where the prime on C'_1 denotes that this is not precisely the same normalisation factor as in Hirata & Seljak (2004), given a numerical value of $5\times10^{-14}(h^2 M_\odot Mpc^{-3})^{-1}$ in Bridle & King (2007).

In fact, this value was normalised to SuperCOSMOS data (Brown *et al.* 2002) at low redshift so we can identify

$$\int \frac{d\chi}{\chi^2}\left(C_1\frac{\rho(z=0)}{D(z=0)}a^3(z=0)\right)^2 n^2(\chi)P_{\delta\delta}(k,z=0) \equiv \int \frac{d\chi}{\chi^2}\left(C'_1\frac{\rho(z_f)}{D(z_f)}a^3(z_f)\right)^2 n^2(\chi)P_{\delta\delta}(k,z_f)$$

$$C'_1 = \frac{C_1\rho(z=0)a^3(z=0)}{D(z=0)\rho(z_f)a^3(z_f)} = C_1\frac{\rho(z=0)}{\rho(z_f)a^3(z_f)} \tag{3.14}$$

This gives us a new source term

$$I(k,z_f) = \frac{C_1 k^2}{4\pi G}\frac{\rho(z=0)}{\rho(z_f)a^2(z_f)}\frac{1}{D(z_f)}\Psi(k,z_f). \tag{3.15}$$

It is interesting to note that this IA formalism is insensitive to the details of galaxy formation, if we can make some common assumptions. If we can assume galaxy formation occurs at high redshift, well within the linear regime, over a relatively short timescale and that the IAs we are interested in can be fully decribed by the LA model, then we do not need to assign a particular redshift of galaxy formation, z_f.

The presence of instrinsic alignments boosts the observed shear correlations. Denoting the galaxy shear and position observables by ϵ and g respectively :

$$C_l^{\epsilon\epsilon} = C_l^{GG} + C_l^{GI} + C_l^{II}, \tag{3.16}$$

$$C_l^{\epsilon g} = C_l^{Gg} + C_l^{Ig}. \tag{3.17}$$

As a benchmark, in figure 3.2.2 we show the gravitational and intrinsic alignment contributions to the shear auto-correlation and and galaxy position-shear cross-correlations for a fiducial ΛCDM cosmology and Euclid-like survey (as described in 3.3.1).

The intrinsic alignment power spectra can be related to the underlying dark matter power spectrum by some intrinsic alignment bias parameters

$$P_{II}^{(obs)} = b_I^2 P_{\delta\delta} \tag{3.18}$$

$$P_{GI}^{(obs)} = b_I r_I P_{\delta\delta} \tag{3.19}$$

(Bernstein 2009). We treat the IA bias parameters in an identical way to the galaxy bias in 3.2.1, described in 3.2.4 below.

3.2.3 Modified Gravity

We are interested in the possibility that a modification to Einstein's GR may be able to explain the late-time expansion of the Universe without the need for the dark energy fluid. There exist a large number of detailed theories which attempt to modify GR in ways that could explain the observed acceleration (Jain & Khoury 2010). These range from $f(R)$ type theories where a function is added to the Einstein-Hilbert action, to extra-dimensional theories, like DGP, which suppose that our 4D spacetime in embedded in a higher dimensional manifold.

Rather than considering a particular, theoretically motivated MG theory, we will attempt to constrain possible deviations from GR. If this approach resulted in a detection of non-GR physics we would not have a ready-made alternative theory of gravity but we would have confirmed that our present theory is incomplete and may even be able to point in the direction of possible alternatives- a valuable outcome considering the huge range of theories currently under discussion.

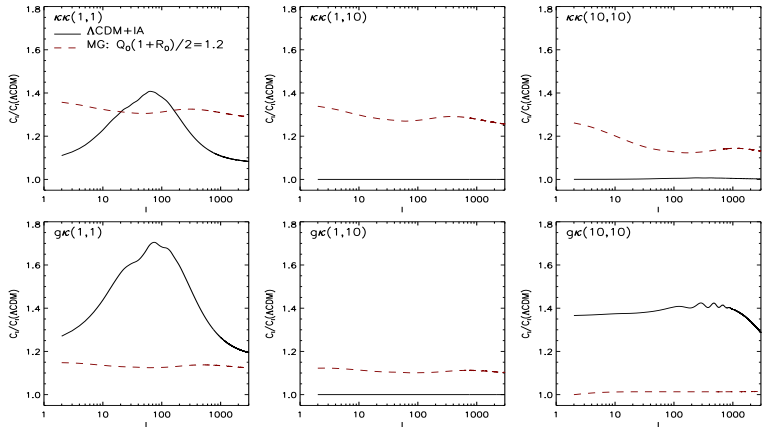

Figure 3.2: Angular power spectra relative to LCDM with no intrinsic alignments, for different redshift bins and with/without intrinsic alignments and modified gravity. Left: Auto-correlation of the lowest redshift bin with itself. Center: Cross-correlation of the lowest redshift bin with the highest redshift bin. Right: Auto-correlation of the highest redshift bin with itself. Upper: Observable shear-shear correlation functions. Lower: Observable shear-position correlation function. Solid: LCDM power spectrum including intrinsic alignments, divided by LCDM power spectrum not including intrinsic alignments. Dashed: Modified gravity power spectrum with $Q_0(1 + R_0)/2 = 1.2$ and no intrinsic alignments, divided by LCDM power spectrum with no intrinsic alignments.

When parameterising deviations from GR, it is useful to consider linear perturbations around the flat FRLW background described in the conformal Newtonian gauge

$$ds^2 = -(1+2\Psi)dt^2 + a^2(1+2\Phi)\delta_{ij}dx^i dx^j \tag{3.20}$$

where a dot denotes a derivative with respect to the cosmic time, t.

The bending of light, and hence WGL, responds to the sum of the metric potentials, $\Phi + \Psi$, while galaxy position information responds only to the Newtonian potential, Ψ.

We consider a phenomenological parameterization for cosmic scale deviations from GR, widely used in the literature (Tereno *et al.* 2011; Zhao *et al.* 2009; Song *et al.* 2010; Bean & Tangmatitham 2010). It employs two functions that modify the Poisson equation and the relationship between the two Newtonian potentials in the conformal Newtonian gauge,

$$k^2\Phi = -4\pi G Q a^2 \sum_i \rho_i \Delta_i \tag{3.21}$$

$$\Psi - R\Phi = -12\pi G Q a^2 \sum_i \rho_i (1+w)\frac{\sigma_i}{k^2}. \tag{3.22}$$

where $Q = R = 1$ for GR. We assume the fluid equations for matter are unaltered. Q effectively boosts the gravitational potential generated by a given mass, so that the effective gravitational constant $G_{eff} = QG$ and is sometimes referred to as the mass screening. $R \neq 1$ alters the relationship between the motion of non-relativistic matter, which is determined by QR and relativistic geodesics, as in lensing, dependent on $Q(1+R)/2$, a value of $R \neq 1$ introduces an effective anisotropic stress, something which is always negligible in pure GR.

We assume that time evolution of the effects of modifications to GR on the growth of inhomogeneities would vary in concert with alterations to the background expansion, and consider modifications of the form, $Q = Q(a) = (Q_0 - 1)a^3$ and $R = R(a) = (R_0 - 1)a^3$. The $\propto a^3$ scaling is invoked to allow a modification to GR which turns on at late times. As such it may be possible to explain late-time cosmic acceleration without contradicting constraints from the CMB and BBN. In future, more general, analysis it would be appropriate to let redshift dependence vary as a (set of) free parameter(s). We omit any scale dependence in our modifications. Two conditions, $Q(a) > 0$ and $R(a) > -1$ are imposed to ensure that overdensities remain gravitationally attractive and that light is bent towards the lens.

3.2.4 Comprehensive Bias Model

We have mentioned the bias and cross-correlation parameters for galaxy and IA terms: b_g, b_I, r_g, r_I. Generally these are free functions of scale and redshift, we choose to follow the parameterisation

of (Joachimi & Bridle 2009; Bridle & King 2007), namely that each of the four terms takes the form

$$b_X(k,\chi) = A_X Q_X(k, z(\chi)) b_X^{\text{base}}(k,\chi), \tag{3.23}$$

where A_X is a amplitude free parameter, $Q_X(k, z(\chi))$ is a 2D grid in k/z-space. The grid has $N_k \times N_z$ nodes, spaced linearly in z and logarithmically in k. Each grid note is a constant parameter, free to vary around the fiducial value of unity. Before it enters the C_l integrand, the grid is interpolated to cover the full range in k/z. Unless otherwise stated, the grid size is assumed to be the fiducial $N_k = N_z = 5$. $b_X^{\text{base}}(k,\chi)$ are the fiducial values of the bias terms which we take to be

$$b_g^{\text{base}}(k,\chi) = 1 \tag{3.24}$$

$$r_g^{\text{base}}(k,\chi) = 1 \tag{3.25}$$

$$b_I^{\text{base}}(k,\chi) = -C_1 \frac{\rho(z=0)}{D(z)} \tag{3.26}$$

$$r_I^{\text{base}}(k,\chi) = 1. \tag{3.27}$$

Note that the b_I term includes our LA model of IAs.

Combining our formalisms for shear, IAs, modified gravity and bias, we can write down the full suite of C_l equations used in this chapter,

$$C_{ij}^{GG}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) \left[Q \frac{1+R}{2} \right]^2 P_{\delta\delta}(k,z) \tag{3.28}$$

$$C_{ij}^{II}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) b_I^2(k,z) Q^2(z_f) R^2(z_f) P_{\delta\delta}(k,z) \tag{3.29}$$

$$C_{ij}^{GI}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) W_j(\chi) Q Q(z_f) R(z_f) (1 + \frac{R}{2}) \\ b_I(k,z) r_I(k,z) \frac{\sqrt{P_{\delta\delta}(k,z_f) P_{\delta\delta}(k,z)}}{D(z_f)} \tag{3.30}$$

$$C_{ij}^{gG}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) n_j(\chi) Q \frac{1+R}{2} b_g(k,z) r_g(k,z) P_{\delta\delta}(k,z) \tag{3.31}$$

$$C_{ij}^{gI}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) Q(z_f) R(z_f) b_g(k,z) b_I(k,z) \tag{3.32}$$

$$r_g(k,z) r_I(k,z) \frac{\sqrt{P_{\delta\delta}(k,z_f) P_{\delta\delta}(k,z)}}{D(z_f)}$$

$$C_{ij}^{gg}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) b_g^2(k,z) P_{\delta\delta}(k,z). \tag{3.33}$$

3.3 Likelihood analysis

3.3.1 Survey specifications

We consider the impact of including IAs on cosmological constraints for a future survey like Euclid (DETF "Stage IV"). We assume the galaxies are distributed

$$n(z) \propto z^2 \exp\left(-\frac{z}{z_0}\right)^{3/2} \tag{3.34}$$

in N_{ph} photometric redshift bins, divided so that they each contain an equal number of galaxies. The photometric redshifts are measured with accuracy $\delta_z(1+z)$. The parameter values assumed to be: area 20000 deg^2, redshift range $0 < z < 3$, projected galaxy density 35arcmin^{-2}, photometric redshift scatter parameter $\delta_z = 0.05$ and intrinsic ellipticity dispersion $\sigma_\gamma = 0.35$ and a median redshift $z_m = 0.9$. z_m is related to z_0 as $z_0 = \frac{z_m}{\sqrt{2}}$.

The noise for each survey is modeled as statistical errors given by

$$N_l^{\epsilon_i\epsilon_j} = \delta_{ij}\frac{\gamma_{rms}^2}{n_j} \tag{3.35}$$

$$N_l^{g_ig_j} = \delta_{ij}\frac{1}{n_j} \tag{3.36}$$

$$N_l^{g_i\epsilon_j} = 0 \tag{3.37}$$

where γ_{rms} is the root mean square ellipticity of galaxies and n_j is number of galaxies per steradian in jth bin. We note that the γ_{rms} (and associated n_g) values for different experiments are defined differently.

We include complementary constraints from temperature (T) and E-mode polarization (E) measurements from a Planck-like CMB survey up to $l = 3000$.

To account for our uncertainty of galaxy biasing effects at small scales we include a cut on scale

$$k_{lin}^{\max}(z) \approx 0.132 z h \mathrm{Mpc}^{-1} \tag{3.38}$$

as per Rassat *et al.* (2007). Instead of cutting the 3D power spectrum in k-space we exclude projected power spectra above a particular angular frequency defined by

$$l_{g(i)}^{\max}(z) = k_{lin}^{\max}(z_{\mathrm{med}}^i)\chi((z_{\mathrm{med}}^i)), \tag{3.39}$$

where we choose the median redshift of a tomographic bin i as the characteristic redshift of that bin. We apply the cut to terms involving the galaxy power spectrum so there is no alteration to $C_l^{\epsilon\epsilon}$. We impose l-cuts on $C_l^{m\epsilon}$, cutting on the bin i where this is the z-bin from which the number density signal stems. For C_l^{mn} we cut on the higher redshift bin.

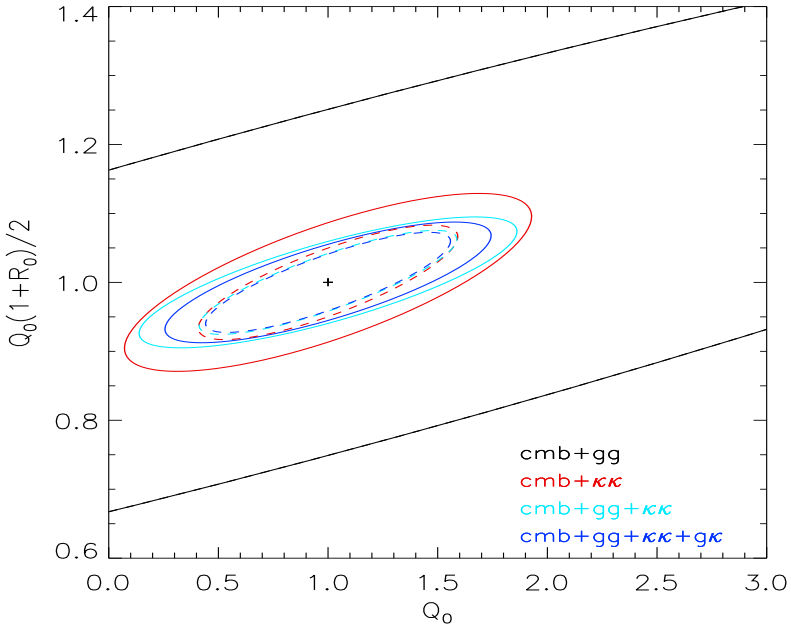

Figure 3.3: 95 per cent confidence constraints on modified gravity parameters showing the effect of adding different datasets. Black solid lines show constraints using just galaxy position-galaxy position power spectra with the CMB. Red solid lines show constraints using just galaxy shear - galaxy shear power spectra with the CMB. Cyan lines show the combination of position-position and shear-shear power spectra with the CMB. Blue solid lines show the complete set of data used in this chapter: CMB, position-position, shear-shear and shear-position power spectra. Dashed lines show the latter three constraints with IAs switched off.

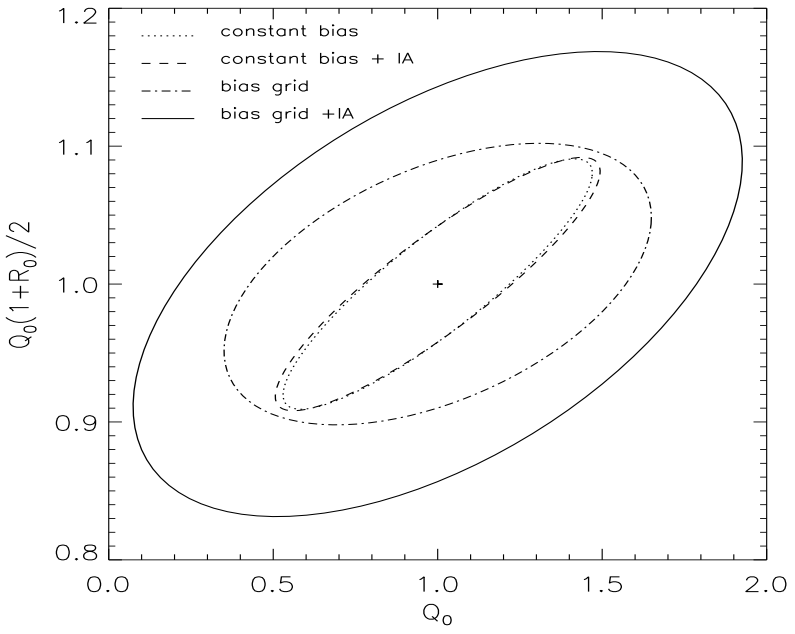

Figure 3.4: 95 per cent confidence constraints on modified gravity parameters for different assumptions about intrinsic alignments. These constraints use shear-shear power spectra, shear-position power spectra and the CMB. The dotted line shows constraints marginalised over just four bias parameters, i.e. assuming the bias functions are constant with scale and redshift, with intrinsic alignments switched off in the predicted power spectra. The dashed line is the same but with intrinsic alignments included in the predicted power spectra. The dot-dashed line marginalises over the four bias functions on a grid in scale and redshift but with intrinsic alignments switched off in the predicted power spectra. The solid line shows results for our recommended scenario in which intrinsic alignments are included and the bias functions are varied on a grid.

3.3.2 Characterizing the likelihood surface

For our analysis we consider constraints on standard cosmological parameters $\Omega_b, \Omega_c, \sigma_8, n_s, \tau, H_0$, dark energy equation of state parameters w_0, w_a, and modified gravity parameters Q_0,and $Q_0(1+R_0)/2$, marginalizing over the galaxy and intrinsic alignment biases b_g, r_g, b_I, r_I. For the bias grid, we investigate the impact of the gridding $N_k = N_z \leq 5$ bins, with logarithimically spaced between $10^{-3}Mpc^{-1} < k < 30Mpc^{-1}$ and $0 < z < z_{max}$.

The large number of parameters, p_i, especially when using the full bias model favors the use of a Fisher matrix approach; for n parameters only $n+1$ power spectrum evaluations are required to estimate the parameter covariance matrix, $C_{ij} = F_{ij}^{-1}$, with,

$$F_{ij} = \frac{\partial t_a}{\partial p_i} \mathrm{Cov}_{ab}^{-1} \frac{\partial t_b}{\partial p_j}. \tag{3.40}$$

where t_a are the set observables $\{C_l^{CMB}, C_l^{gg}, C_l^{g\epsilon}, C_l^{\epsilon\epsilon}\}$ across all redshift bin combinations. We have modified the publicly available CosmoMC and CAMB (Lewis & Bridle 2002) codes to calculate the Fisher Matrix and the correlation functions for the future survey specifications, in light of the IA and modified gravity models. We ignore covariances between the CMB and large-scale structure here and treat the Fisher matrix from the CMB as independent from the Fisher matrix from the large-scale structure observables. We include the covariances between the large-scale structure observables as specified in Joachimi & Bridle (2010).

3.4 Analysis

We illustrate the effect of intrinsic alignments on cosmic shear power spectra in Fig. 3.2. The intrinsic alignment contribution is larger at lower redshifts (left panels) and decreases to higher redshifts (right panels). This happens even though the source three-dimensional power spectra are constant with redshift, due to the inverse square comoving distance in the angular power spectrum equation. Conversely the lensing power spectra increase with redshift because there is more material to integrate over and also because lensing increases with the distances in the system even for a fixed mass lens.

The upper panels in Fig. 3.2 show the usual cosmic shear power spectra in which the shear of one galaxy is correlated with the shear of a second galaxy. If the galaxies are on average pointing in the same direction then the power spectrum is positive. This is the case for the usual lensing signal and for the II intrinsic alignment term. For the GI term one galaxy is lensed tangentially around a dark clump and one galaxy is stretched during formation by the dark clump and so the correlation

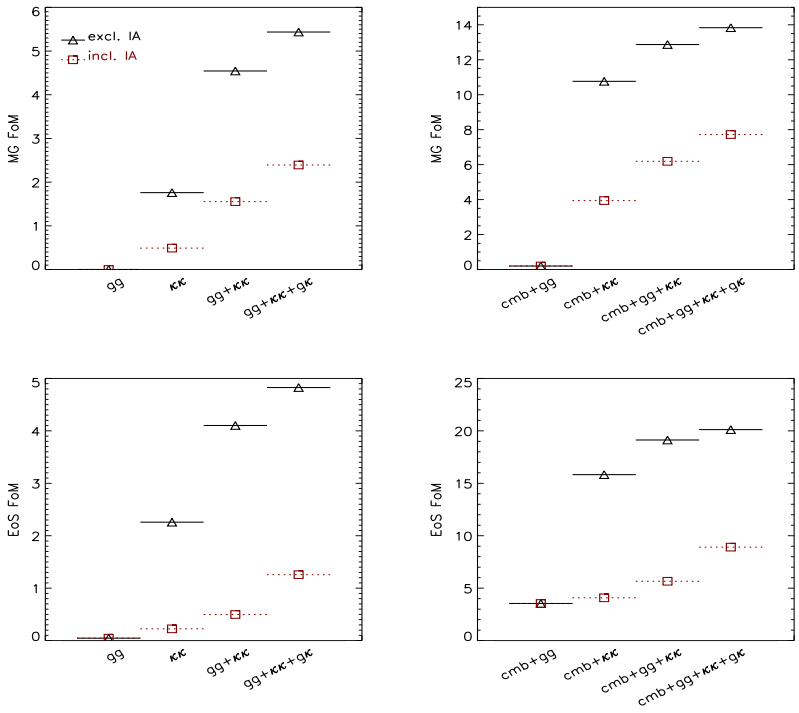

Figure 3.5: Figures of merit for different data combinations. Upper panels: modified gravity figure of merit. Lower panels: dark energy figure of merit. Left panels: without CMB. Right panels: with CMB. Solid lines are with intrinsic alignments switched off. Dotted lines include intrinsic alignments. Individual columns are explained in the text.

is negative. Because we freeze the intrinsic alignment model at the redshift of formation, the II term has no non-linear contributions on small scales, whereas the GI term has some non-linear contributions due to the gravitational shearing part, and the usual cosmic shear power spectrum has the full non-linear contribution. This is most apparent in the left top panel because for the furthest redshift bin (top right) the non-linear physical scale has moved close to the maximum angular scale we consider.

The lower panels show the extra observable, the shear-position correlation function, in which the shear of one galaxy is correlated with the position of the second galaxy. If the first galaxy is pointing at the position of the second galaxy then this correlation function is positive, by our

choice of sign convention. The solid line in the lower panels arises from the gravitational lensing of more distant galaxies by dark matter hosting closer galaxies. For simplicity here we do not include the small contributions from the galaxy haloes themselves (Bridle & Abdalla 2005). This term is largest when one redshift bin is at twice the distance of the closer redshift bin and it is negative because lensing causes tangential alignments around dark matter clumps. The dashed line shows the contribution from the intrinsic alignment of one galaxy relative to the local dark matter which also hosts other nearby galaxies. This signal is largest when both galaxies are that the same redshift and is positive in the case that galaxies are tidally stretched along the strong curvature in the direction of a dark matter clump.

Both the modified gravity parameters and IA parameters can modify the predicted shear and position-shear correlations in comparison to the bare cosmic shear prediction. In figure 3.2 we show that through using tomographic redshift slices and the combination of shear-shear and shear-position data we might hope to break the degeneracy between the two. The relative amplitudes of the IA and MG curves in the top left plot is clearly very different from that in the top centre or top right. The same is true for the bottom left versus the bottom centre and bottom right. It is this difference in amplitude and variation with k, dependent on which combination of tomographic bins we are looking at, which suggests we may be able to break the degeneracy between IAs and MG parameters. On top of this the fact that the top and bottom panels show different changes in behaviour, especially for the left hand panels, is encouraging. The combination of different probe combinations may aid the tomography in breaking degeneracies.

It is often the case that in projecting cosmological constraints from large scale structure surveys a simple constant bias model is employed to relate the observed galaxy power spectrum to the underlying matter correlations. Therefore in figure 3.3 we show the impact of extending beyond a simple constant bias model. The constraints from galaxy clustering (position-position power spectra) alone are, as expected, very tight on using a constant bias model and are completely degraded on marginalising over a galaxy bias model with many free parameters as a function of scale and redshift. We focus in figure 3.3 on constraints from CMB, shear-shear and shear-position data. The constraints are tightest when using the constant bias model and when intrinsic alignments are switched off (dotted line). When intrinsic alignments are included this weakens the constraints even for the constant bias model because some of the cosmological constraining power of cosmic shear is lost in constraining the intrinsic alignment model. It has been shown that cosmic shear alone constraints are significantly degraded on including intrinsic alignments and marginalising over a flexible grid (Bridle & King 2007, Joachimi & Bridle 2010) but we see

that once CMB and shear-position correlations are included with cosmic shear this effect is less severe. The solid line includes intrinsic alignments and a very flexible model and is about a factor of 4 larger in area than the most optimistic scenario (dotted line). This is partly due to the effect of marginalising the galaxy bias within the shear-position power spectrum (dot-dashed line compared to the dotted line) and partly due to marginalisation over the intrinsic alignment contribution (solid line compared to the dot-dashed line).

In figure 3.4 we summarize the constraints with and without IAs for CMB plus galaxy, lensing and galaxy lensing cross correlations separately. The use of a grid bias model leads to the galaxy-galaxy correlations providing little constraining power on the non-bias parameters in the model. When lensing and galaxy position data are added in tandem, however, in combination they are able to provide improved constraints over and above the lensing data alone. When IAs are included the combined constraints are noticeably weaker than the constraints from predicted by pure shear-shear measurements when IAs are neglected.

We define a figure of merit (FoM) for inhomogeneous growth using the modified gravity parameters Q_0 and $Q_0(1 + R_0)/2$, and the background expansion, using the equation of state, w_0 and w_a. In each case the FoM is the inverse area for the 95% confidence ellipse in the 2D likelihood space, with all other parameters are marginalized over: FoM=$1/6.17\pi ab$, where a and b are the semi-major and semi-minor axes of the ellipse describing the covariance matrix for the pair of parameters.

In figure 3.5 we quantify the impact of IAs on the constraints from different combinations of data. When the constraints from large scale structure measurements alone are considered the inclusion of IAs results in a dramatic reduction in both the modified gravity and equation of state figures of merit (up to $\sim$80%). Only by including the cross correlations between galaxy position and lensing observations can we recover a FoM comparable to that predicted by lensing alone without IAs. When we include prospective CMB data the impact is reduced but is still marked, with a near 50% reduction in both FoMs for CMB+lensing and galaxy position data combined. With the inclusion of the CMB data, the combination of all probes, $CMB+\epsilon\epsilon+n\epsilon+nn$, including IAs no longer recovers the same constraining power as shear-shear, $CMB+\epsilon\epsilon$, alone without IAs. Even so, there is still roughly a doubling of constraining power, when we include IAs, between the shear-shear only case and the case where we have added position-position and position-shear correlations.

In figure 3.6 we highlight the impact the bias grid parameterization has on the degradation in the figure of merit. For the modified gravity parameters, even just a two bin bias model gives a

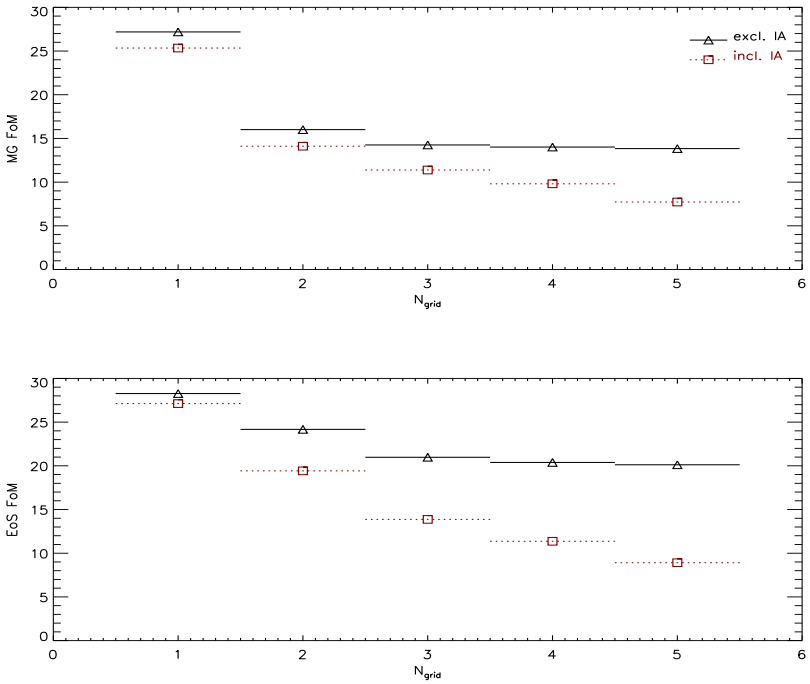

Figure 3.6: Dependence of figures of merit on the number of parameters used to specify the bias functions. We set the number of grid parameters in the scale and redshift directions equal to the number on the x-axis. Solid lines are with intrinsic alignments switched off. Dotted lines include intrinsic alignments.

noticeable reduction in the FoM in comparison to the commonly used constant bias case. While if IAs are excluded additional bins in the grid model have a small additional effect, the same is not true when the IAs are included, and, perhaps unsurprisingly, the disparity between predictions between with and without IAs becomes more marked the finer gridding is used. It may be that the particular redshift dependencies of the shear-shear and IA components of the lensing signal contribute to the degradation in FoM for the IA cases. More grid nodes confer more uncertainty on our understanding of how the cosmic shear signal varies with redshift which has a greater impact once we include IAs. It is worth noting that the decrease from 1 to 2 bins is much more severe for the MG FoM than for the dark energy FoM, suggesting that MG constraints are more sensitive to

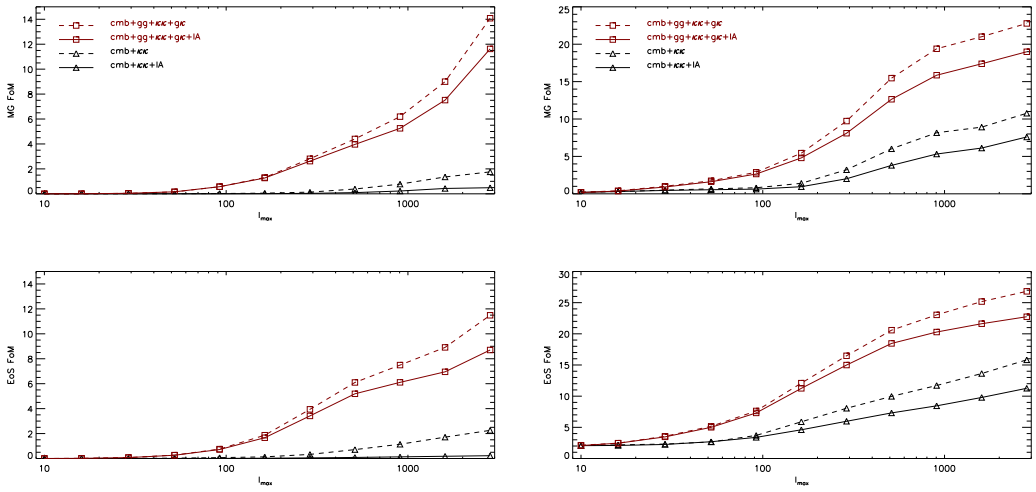

Figure 3.7: Figures of merit as a function of maximum multipole used in the shear-shear power spectrum analysis. Upper panels: modified gravity figure of merit. Lower panels: dark energy equation of state figure of merit. Left panels: no CMB. Right panels: with CMB. Black lines show results using just shear-shear power spectra and CMB data. Red lines show the full dataset used in this chapter. Dashed lines have intrinsic alignments switched off in the predicted power spectra. Solid lines include intrinsic alignments.

a small level of redshift uncertainty. Interestingly the decrease between 2 and 5 bins (inc. IAs) is more pronounced for dark energy than for MG. This suggests that the damage is done to the MG constraint by the first few bins in k/z.

Figure 3.7 shows the effect of truncating the cosmic shear power spectra at smaller maximum spherical harmonic number $\ell_{\rm max}$. Note that as discussed in the previous section, we already truncate the position-position and position-shear power spectra at lower $\ell_{\rm max}$. This is important because it becomes increasingly difficult to make theoretical predictions small scales (high ℓ). It is a problem for conventional cosmic shear analyses because the dark matter clustering must be predicted on small scales (see e.g. Huterer et al. 2008 for a discussion). It is an even bigger problem if intrinsic alignments are included because the intrinsic alignment bias and correlation coefficient could become much more complicated functions of scale and redshift at small scales, and might not be sufficiently described by the grid of parameters we have used. This could provide a reason for cutting off cosmic shear analyses at some smaller ℓ even below those described well by dark

matter predictions.

We see from figure 3.7 that if CMB information is not included then the amount of information obtained from the smallest scales can increase dramatically. This is particularly true for the modified gravity figure of merit (upper left panel), which increases by a factor of two between a conservative $\ell = 1000$ and our fiducial $\ell = 3000$ (with all large scale structure data and intrinsic alignments, solid line). This corresponds to changing the survey area by a factor of 2 (e.g. see Amara & Refregier 2006 and discussion in chapter 4). The figure of merit for dark energy (lower left panel) is less sensitive to the exact maximum ℓ but still increases by 50% between $\ell = 1000$ and $\ell = 3000$. Fortunately with the inclusion of CMB data we can afford to take a safer cut on $\ell_{\rm max}$ whether we are interested in constraining modified gravity (upper right panel) or dark energy (lower right panel). This conclusion is strongest when considering the full set of data used in this chapter and including intrinsic alignments.

3.5 Conclusions

It has been proposed that intrinsic alignments can be removed using shear-position cross-correlation information (Zhang 2009, Joachimi & Bridle 2010) and also that modified gravity can be constrained well using this same information (Zhang 2010). In this chapter we investigate whether both can be constrained by the same information.

We have illustrated that intrinsic alignments and modified gravity affect the shear observables in different ways, with intrinsic alignments dominating at low redshift and modified gravity contributing at a range of redshifts. Future surveys plan to divide up the signal into a number of tomographic redshift bins (Hu 1999), or otherwise use the three-dimensional information (e.g. Heavens 2007, Kitching et al. 2009). This allows for additional information to be used in the cross-correlation between redshift bins. Here we also see a big difference in the type of signal from intrinsic alignments and modified gravity. Intrinsic alignments are a localised effect and therefore again they have a different signature. These differences extend into the shear-position correlation functions, which again have a more localised effect for intrinsic alignments but this time it is significant at all redshifts in the auto spectra for a particular redshift bin.

We reproduced earlier results (Joachimi & Bridle 2010) which found that using all large-scale structure two point functions and marginalising over intrinsic alignments can give a comparable dark energy figure of merit to a conventional cosmic shear analysis using shear-shear power spectra alone and ignoring intrinsic alignments. The gap widens however when CMB is included. Since

Planck has already taken much data then it now makes sense to include this CMB information in any cosmological predictions. Excluding IA from dark energy figure of merit calculations gives the impression that shear-shear power spectra alone can make a large increase in figure of merit, but that additional large scale structure two point functions are not particularly helpful. However when intrinsic alignments are included marginalising over flexible bias functions of scale and redshift then this conclusion is reversed. Shear-shear power spectra alone do not make a big improvement on the figure of merit from the CMB alone, but when the shear-position power spectra are included the figure of merit can be doubled relative to the CMB.

We repeated this investigation but instead considering a modified gravity figure of merit. We use a simple phenomenological form for the modifications, using the ratio of metric potentials as one free parameter, and the prefactor in the Poisson equation as the other. We allow a flexible expansion history parameterised by the dark energy equation of state parameters, to mimic the potential background behaviour of a new theory of gravity. The CMB alone therefore has no constraints on the modified gravity figure of merit.

Constraints from large-scale structure alone show the same trends as for the modified gravity figure of merit, with increasing constraining power as more two point functions are added. This is true with or without intrinsic alignments, however this time the combination of all the large-scale structure data marginalised over intrinsic alignments is now better than the conventional cosmic shear analysis. With the CMB included we see that the large-scale structure measurements can make significant constraints on modified gravity parameters even in the presence of intrinsic alignments. These constraints are tightened when the shear-position correlation function is included, showing that this extra observable has the capacity to tighten constraints on intrinsic alignments and modified gravity simultaneously.

We have also shown how our results are robust to different numbers of free parameters in the galaxy clustering and intrinsic alignment bias models, and investigated the dependence on the maximum spherical harmonic number included.

If the intrinsic alignments are included without any bias, they actually form a unique signal to add to the traditional shear and so add constraining power, while as we go to constant and gridded bias models, the contraints clearly deteriorate. This stresses the importance of developing a physically motivated better understood model for how galaxies align themselves in a shared potential, as then the systematic could in theory be turned into an observable.

It should be noted that the treatment of IAs in this and the following chapter is in some respects more sophisticated than that in chapter 2 but somewhat less realistic in some others. The more

advanced treatment of scale- and redshift-dependent nuisance parameters is a great improvement of the previous 2-parameter model. On the other hand, for reasons of clarity, we omit the "galaxy property" parameters of f_r and β in chapters 3 & 4. Nor do we employ the more physically motivated halo model. While this helps to foreground the impact of the effects we are interested in- IAs, galaxy position information and survey strategy - it would be important in future forecasts and analysis of real data to include all relevant information about the galaxy properties sampled.

CHAPTER 4

MEASURING MODIFIED GRAVITY WITH FUTURE COSMIC SHEAR SURVEYS INCLUDING INTRINSIC ALIGNMENTS

4.1 Introduction

The dawn of "precision cosmology" was heralded by the results of surveys which, for the first time, produced data of sufficient quantity and quality that our cosmological probes could begin to accurately measure some of the fundamental properties of the Universe. The result has been the ΛCDM concordance cosmology, a description of the Universe compatible with the joint constraints from many sources of cosmological information. The next decades will see a step-change in our ability to measure cosmological parameters as new surveys produce orders of magnitude more data than has previously been available. This will allow us to test the standard cosmological model as never before.

The standard ΛCDM model describes a Universe made up of ∼75% Dark Energy, a smooth, negative pressure fluid, ∼20% Dark Matter, collisionless massive particles which interact solely via gravity, and just ∼5% baryons making up the visible mass of the Universe (Dunkley *et al.* 2009). The standard model also assumes gravity to be described by Einstein's General Relativity. Testing this theory of gravity on cosmic scales will be one of the most interesting opportunities afforded by upcoming survey data.

The simultaneous observation of accelerating cosmic expansion by a number of groups studying type Ia supernovae was one of the most important discoveries in recent times. A Universe

governed by GR and populated by standard gravitating matter cannot explain the observed acceleration. The dark energy fluid was proposed to solve this paradox, reviving the idea of a cosmological constant which Einstein had originally included in his GR field equations and subsequently discarded. Much effort has been devoted to characterising the nature of dark energy- particularly attempts to discriminate between a pure cosmological constant and a dynamic scalar potential with time-varying equation of state w(a) (Albrecht *et al.* 2006). In all its forms, dark energy poses problems of fine-tuning for which we have, as yet, no physical motivation.

This being the case, it makes sense to also investigate an alternative explanation for cosmic acceleration. Rather than invoke dark energy we may suggest that our standard theory of gravity is incomplete and that a correct theory of gravity would explain cosmic acceleration at late times and large scales in a purely CDM universe. There are a large number of theoretically motivated modified theories of gravity, see Jain & Khoury (2010) for a review. We follow I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011) in attempting to measure general deviations from GR via two "trigger parameters" Q and R.

Weak Gravitational lensing is a particularly useful probe of gravity because it is sensitive to $\Phi + \Psi$, the sum of the metric potentials. Through the modified growth of structure and geometry via the lensing integral cosmic shear can constrain both trigger parameters. In contrast probes such as galaxy redshift surveys and galaxy peculiar velocities depend only on the Newtonian potential, Ψ. WGL constraints also have the benefit that they probe the dark matter distribution directly, avoiding the impact of galaxy biasing. Joint constraints from combining multiple probes can help break degeneracies between the effects of the metric potentials, producing the tightest constraints.

Cosmic shear is the name given to weak gravitational lensing in random patches of the sky and was first detected a decade ago (Kaiser et al. 2000, Wittman et al. 2000, Refregier et al. 2000, van Waerbeke et al. 2000). The latest constraints on cosmology from come from the Hubble Space Telescope COSMOS survey (Schrabback et al. 2010, Massey et al. 2009), and the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) (Fu et al. 2006). In addition the 100 square degree survey (Benjamin et al 2007) combines data from several smaller surveys (Hoekstra *et al.* 2006; Hetterscheidt *et al.* 2007; Le Fèvre *et al.* 2004; Hoekstra *et al.* 2002).

Cosmic shear has been identified as the method with the most potential to uncover the nature of dark energy (Albrecht et al. DEFT reports, Peacock et al. ESA/ESO WG report) and therefore a number of surveys are planned with a major cosmic shear component. Upcoming "Stage III" projects include Kilo-degree Survey (KIDS) on the Very Large Telescope (VLT) Survey Telescope (VST), the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS) project, the

Subaru Meaurement of Images and Redshifts (SuMIRe) survey using HyperSUPRIMECam, the Dark Energy Survey (DES) on the Blanco Telescope. More ambitious "Stage IV" projects are the Large Synoptic Survey Telescope (LSST) ground-based project and in space the proposed European Space Agency mission Euclid and the NASA proposed Wide-Field Infrared Survey Telescope (WFIRST).

Greater accuracy demands better treatment of systematic effects. The main systematic in cosmic shear studies is the contamination by the intrinsic alignment of galaxy shapes. In the previous chapter we showed the strong degeneracy between the IA contamination and our MG parameters. Including a realistic IA model and marginalising significantly degrades the constraining power of a cosmic shear survey. What is more encouraging is that the use of position-position correlations nn and particularly position-shear $n\epsilon$ cross-correlations can go some way towards mitigating the impact of IAs. This data is already collected by a standard WGL survey. In the presence of IAs, constraints from $\epsilon\epsilon + n\epsilon + nn$ are roughly twice as strong as those from $\epsilon\epsilon$ alone. Until now the discussion of optimal survey parameters has generally assumed that $\epsilon\epsilon$ correlations alone will be included, without IAs. In this chapter we investigate the impact of survey strategy and design on different combinations of probes, with and without IAs.

The observing strategy of a given survey is central to the type of data and the quality of results that survey will produce. Different survey geometries with the same instrument will produce different statistical errors due to a different balance between survey characteristics, the most important for cosmic shear being survey area, A, galaxy number density on the sky, n_g and median redshift of the galaxy distribution, z_m. Other properties such as required redshift accuracy are important guidelines for instrument designers and those preparing analysis pipelines and follow-up or calibration studies (Amara & Refregier 2006). While more area and increased number density will always be desirable, it is important to keep firm goals in mind when producing desiderata for future surveys which will always be limited by technology and finite observing time. In particular learning about dark energy and modified gravity may benefit from different survey strategies and call for the prioritisation of different properties.

In this chapter we take a fiducial model for a stage-IV cosmic shear survey and study the impact of changing some of the important survey parameters, examing the impact of the change on the survey's ability to constrain dark energy and modified gravity. We try to include restrictions due to finite total observing time and/or finite budgets which will always limit the scope of real world surveys.

The cha[ter is organised as follows: in section 2, we summarise our standard cosmology, fidu-

cial surveys and models for deviations from GR, Intrinsic Alignments and bias parameterisation, as well as our figures of merit for DE and MG. In section 3 we present the effect of varying survey area and the impact of finite survey time on overall strategy and the ability to constrain MG. The importance of photometric redshift accuracy is addressed in section 4 and the utility of spectroscopic redshifts in section 5. Conclusions are made in section 6.

4.2 Cosmological Set-Up

The chapter deals with the impact of various aspects of survey strategy and redshift quality on the power of cosmic shear and galaxy position information to constrain dark energy and deviations from General Relativity. In 4.2.1 we summarise some basic cosmology we use throughout the chapter. 4.2.2 summarises our cosmic shear formalism, 4.2.3 extends it to include intrinsic alignments and 4.2.4 adds galaxy position auto- and cross-correlations and introduces a coherent biasing formalism for IAs and galaxy bias. 4.2.5 reviews the MG parameterisation we use and how it enters our angular power spectrum integrals. This completes our review of the formalism introduced in I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011), whose main results are summarised in 4.2.6. 4.2.7 describes the fiducial survey parameters we use in the following sections.

4.2.1 General Cosmology

Throughout this chapter we assume a flat ΛCDM cosmological model with fiducial parameter values equal to the WMAP7 best-fit values: $\Omega_m = 0.262$, $\Omega_b = 0.0445$, $w_0 = -1$, $w_a = 0$, $\sigma_8 = 0.802$, $h = 0.714$, $n_s = 0.969$, $n_{run} = 0$, $\Omega_\nu = 0$. The linear matter power spectrum is given by the fitting formula of Eisenstein & Hu (1998), with nonlinear corrections from Smith *et al.* (2003).

When we treat MG we assume that the background expansion of the Universe is consistent with our fiducial ΛCDM model and deviations due to non-GR physics enter through the growth of structure from initial perturbations. Unless otherwise specified all plots include CMB priors from a Planck-type survey.

We calculate constraints on cosmological parameters using the Fisher Matrix formalism

$$F_{\mu\nu} = \sum_{m,n}^{N_d} \sum_{l}^{N_l^{max}(m,n)} \frac{\partial \mathscr{D}_m(l)}{\partial p_\mu} Cov^{-1}_{mn}(l) \frac{\partial \mathscr{D}_n(l)}{\partial p_\nu} \tag{4.1}$$

where m, n are combinations of tomographic redshift bins, $\mathscr{D}_m(l)$ is the data vector, in our case some combination of $\hat{C}_l^{\epsilon\epsilon}$, $\hat{C}_l^{m\epsilon}$ and $\hat{C}_l^{mn}$ where $\hat{C}_l$ is the sum of the angular power spectrum

and the appropriate noise power spectrum. $Cov_{mn}(l)$ is the covariance matrix, p_μ is one of the parameters we vary and N_d is the number of independent combinations of tomogrpahic bins. Unless otherwise stated the Fisher matrix is composed of the following cosmological parameters: $p = \{\Omega_m, w_0, w_a, h, \sigma_8, \Omega_b, n_s\}$.

Using the Fisher matrix formalism the lower limit on the minimum variance bound on the error of a parameter p_μ marginalised over all other parameters of interest, is given by $\sigma(p_\mu) = \sqrt{(F^{-1})_{\mu\mu}}$. We quote results in terms of the DETF FoM for Dark Energy

$$FoM_{DE} = \frac{1}{4\sqrt{\det\left(F_{GR}^{-1}\right)_{w_0,w_a}}} \tag{4.2}$$

where $(F_{GR}^{-1})_{w_0,w_a}$ means that the GR cosmological FM is inverted (MG parameters fixed), then the rows/cloumns associated with w_0 and w_a, from the parameterisation of the dark energy equation of state $w = w_0 + (1-a)w_a$, are extracted. The inversion marginalises over cosmological and nuisance parameters only i.e. MG parameters are excluded and fixed at their GR values. This is effectively proportional to the inverse of the area of the constraint contour in $w_0 - w_a$ space. By analogy we define a MG FoM as

$$FoM_{MG} = \frac{1}{4\sqrt{\det\left(F^{-1}\right)_{Q_0,Q_0(1+\frac{R_0}{2})}}} \tag{4.3}$$

where we have inverted the full FM and marginalised over all parameters (including MG parameters) then extracted the rows/columns associated with Q_0 and $Q_0(1+\frac{R_0}{2})$.

4.2.2 Cosmic Shear

Cosmic shear is the shape-distortion induced in the image of a distant galaxy due to the bending of its light by gravity as it passes massive structure in the Universe. We define the cosmic shear angular power spectrum under the Limber approximation as

$$C_{ij}^{GG}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) P_{\delta\delta}(k, z) \tag{4.4}$$

where $P_{\delta\delta}(k, z)$ is the three dimensional matter power spectrum, χ is the comoving distance in units of h^{-1}Mpc and $W(\chi)$ is the lensing efficiency function

$$W_i(\chi) = \frac{4\pi G}{c^2}\rho(z)a^2(z)\chi \int d\chi' n_i(\chi') \frac{\chi' - \chi}{\chi'} \tag{4.5}$$

4.2.3 Intrinsic Alignments

The intrinsic alignment (IA) of galaxy ellipticities is a prime contaminant to the measured cosmic shear signal. A naive approach to cosmic shear assumes that galaxy's intrinsic ellipticity are randomly distributed on the sky so, when we average over observed ellipticity in a small patch, intrinsic ellipticity cancels and we are left with the induced shear.

Unfortunately this assumption is invalid because galaxy ellipticities are aligned due to two effects arising from the same physical intrinsic alignment origin. Physically close galaxies tend to align with the local gravitational tidal field and so are positively correlated with each other, this is the Intrinsic-Intrinsic (II) alignment. Background galaxies can have their light lensed by foreground gravitational fields which align the intrinsic ellipticity of foreground galaxies. This induces an anti-correlation and is the gravitational-intrinsic (GI) alignment.

We follow the procedure of I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011) and implement IAs using an updated version of the corrected LA model of Hirata & Seljak (2004), incorporating their correction of *erratum 2010* and, assuming that all IA physics occurs at a high-z epoch of galaxy formation, we can write an IA source term as

$$I(k,z) = \frac{C_1 k^2}{4\pi G}\frac{a(z)}{D(z)}\Psi(k,z) \tag{4.6}$$

where G is Newton's constant, $D(z)$ is the linear growth function, $\Psi(k,z)$ is the Newtonian potential and $C_1 = 5 \times 10^{-14}(h^2 M_\odot Mpc^{-3})^{-1}$ is the amplitude of the IA term, normalised to redshift zero (Bridle & King 2007). This reference extracted a value for C_1 from Mandelbaum *et al.* (2005), at the time errors on C_1 were very large but we believe it is well suited for treatment as the fiducial value of a nuisance parameter to be marginalised over. Future surveys will produce tighter constraints on C_1 and our IA model more generally. 4.6 can be used to generate projected angular power spectra for the Intrinsic-Intrinsic (II) and the Gravitational-Intrinsic (GI) correlations, see table 4.1. The window function associated with 4.6 is the galaxy redshift distribution, $n(z)$.

The total observed lensing signal is then the sum of the cosmic shear and the IA terms $C_l^{\epsilon\epsilon} = C_l^{GG} + C_l^{II} + C_l^{GI}$.

4.2.4 Galaxy Position Data

A cosmic shear survey contains galaxy position information (angular position on the sky and redshift) as well as measurements of galaxy shear. Joachimi & Bridle (2010) provide a formalism

for including this additional information in our cosmological parameter constraints and show how this extra information can serve to partially overcome the impact of IAs.

The extra observables we use are galaxy position-position power spectra and the position-shear cross-spectra, defined analogously to the shear-shear power spectrum:

$$C_l^{nn} = C_l^{gg} \tag{4.7}$$

$$C_l^{n\epsilon} = C_l^{gI} + C_l^{gG}. \tag{4.8}$$

We now assign nuisance parameters to each component C_l in a self-consistent way as explained in I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011). The four bias parameters are b_g,b_I,r_g and r_I. b_g and b_I model bias from galaxy position and IAs, and their cross-correlations are modelled by r_g and r_I. Each has one free amplitude parameter and a $n_k \times n_z$ grid of free parameters interpolated over the full range in k-,z-space. Unless otherwise stated the grid size is set to the fiducial value of $n_k = n_z = 5$. The nuisance parameters multiply linearly into the angular power spectra integrands and are included as:

$$\begin{array}{llll} C_l^{GG}: & --, & C_l^{gG}: & b_g r_g \\ C_l^{II}: & b_I b_I, & C_l^{gI}: & b_g b_I r_g r_I \\ C_l^{GI}: & b_I r_I, & C_l^{gg}: & b_g b_g \end{array} \tag{4.9}$$

We have ignored the effect of lensing magnification and followed I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011) and Joachimi & Bridle (2009) in cutting any redshift bin combination ij including galaxy position information at $k_{lin}^{max}(z_{med}^{(i)})\chi(z_{med}^{(i)})$ to account for uncertainties in the galaxy bias model at small scales.

4.2.5 Modified Gravity

There are a large number of modifications or extensions of Einstein's general theory of relativity which come under the general heading of Modified Gravity theories. These can be motivated by the presence of extra dimensions, as in DGP, or extra degrees of freedom compared to the GR action equation, as in f(R), TeVeS etc.

In this chapter we do not assume a particular modified theory of gravity but rather concentrate on "trigger parameters" whose deviation from their GR values would indicate the presence of some physics beyond that in the standard GR picture.

In the conformal newtonian gauge the metric for a flat FRW spacetime is written

$$ds^2 = -a(\tau)^2 \left[1 + 2\Psi(x,t)\right] d\tau^2 + a(\tau)^2 \left[1 - 2\Phi(x,t)\right] dx^2 \tag{4.10}$$

where Ψ and Φ are the scalar potentials which describe perturbations to the time- and space-parts of the metric respectively.

We follow the approach of I. Lazslo, D. Kirk, R. Bean and S. Bridle (2011) and parameterise deviations from GR through two parameters, Q and R. One alters the way the Newtonian potential responds to mass via the Poisson equation,

$$k^2\Psi(x,t) = -4\pi G Q \rho a^2 \delta \tag{4.11}$$

and the other modifies the ratio of the metric potentials

$$\Psi(x,t) = R\Phi(x,t) \tag{4.12}$$

Q and R are assumed to be scale-independent and vary with redshift as

$$Q = (Q_0 - 1)a^s \tag{4.13}$$

$$R = (R_0 - 1)a^s \tag{4.14}$$

where a is the scale-factor and Q_0,R_0 are the free parameters we vary for MG, and $s = 3$. We are interested in MG theories which explain the accelerating expansion of the Universe observed at late times. $s = 3$ allows any modification to "turn-on" at late times and avoids violating early Universe constraints from the CMB & BBN.

As well as modifying the linear growth equation, Q and R enter the angular power spectrum integrals via their response to the metric potentials through the linear growth function.

The full projected angular power spectra, including IAs, MG and bias parameters, is summarised in Table 4.1.

4.2.6 Summary of Chapter 3 results

Our previous chapter (I. Lazslo, D. Kirk, R. Bean and S. Bridle 2011) investigated the degeneracy between IAs and MG parameters. It was found that the inclusion of a realistic IA model reduced the FoM (dark energy or MG) for a typical stage-IV comic shear survey $\sim$65%. The effect could be mitigated by the inclusion of galaxy position data. The constraining power of $\epsilon\epsilon + n\epsilon + nn$, including IAs, is roughly double that of $\epsilon\epsilon$, including IAs, alone.

We showed that our results were robust to different numbers of free parameters in the galaxy clustering and IA bias models.

Table 4.1: Summary of the projected angular power spectra considered in this work.

measured correlation	2D PS
shear	$C_{ij}^{GG}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) \left[Q(z) \frac{1+R(z)}{2} \right]^2 P_{\delta\delta}(k,z)$
intrinsic-shear	$C_{ij}^{GI}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) n_j(\chi) Q(z) Q(z_f) R(z_f) (1 + \frac{R(z)}{2}) b_I(k,z) r_I(k,z) \frac{\sqrt{P_{\delta\delta}(k,z_f) P_{\delta\delta}(k,z)}}{D(z_f)}$
intrinsic	$C_{ij}^{II}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) b_I^2(k,z) Q^2(z_f) R^2(z_f) P_{\delta\delta}(k,z)$
galaxy clustering	$C_{ij}^{gg}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) b_g^2(k,z) P_{\delta\delta}(k,z)$
clustering-shear	$C_{ij}^{gG}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) W_j(\chi) Q(z) \frac{1+R(z)}{2} b_g(k,z) r_g(k,z) P_{\delta\delta}(k,z)$
clustering-intrinsic	$C_{ij}^{gI}(l) = \int \frac{d\chi}{\chi^2} n_i(\chi) n_j(\chi) Q(z_f) R(z_f) b_g(k,z) b_I(k,z) r_g(k,z) r_I(k,z) \frac{\sqrt{P_{\delta\delta}(k,z_f) P_{\delta\delta}(k,z)}}{D(z_f)}$
galaxy ellipticity (observable)	$C_{ij}^{\epsilon\epsilon} = C_{ij}^{GG} + C_{ij}^{II} + C_{ij}^{GI}$
galaxy number density (observable)	$C_{ij}^{nn} = C_{ij}^{gg}$
number density-ellipticity (observable)	$C_{ij}^{n\epsilon} = C_{ij}^{gI} + C_{ij}^{gG}$

4.2.7 General Survey Parameters

In this chapter we vary certain survey parameters around a fiducial cosmic shear survey set-up, corresponding to Euclid, a proposed stage-IV mission which would launch by the end of the decade. This survey has an area = 20000deg^2, number density of galaxies projected onto the sky, $n_g = 35$, and galaxy intrinsic ellipticity dispersion, $\sigma_\gamma = 0.35$, a photometric redshift scatter parameter $\delta_z = 0.05$, with a fraction of catastrophic outliers, $f_{\rm cat} = 0$. We analyse the results using cosmic shear tomography with 10 tomographic redshift bins of equal number density.

The redshift distribution of galaxies, $n(z)$ is assumed to be given by a Smail-type distribution (Smail *et al.* 1994)

$$n(z) \propto z^\alpha exp\left(-\frac{z}{z_0}\right)^{(\beta)} \tag{4.15}$$

with $\alpha = 2$,$\beta = 1.5$ and $z_0 = z_m/1.412$, z_m = 0.9, see e.g. Amara & Refregier (2007) for these fiducial values.

4.3 Survey area

In this Section we calculate how cosmological parameter constraints depend on survey area, investigating the difference between dark energy and modified gravity parameters and the impact of intrinsic alignments.

We first consider the simple case where all survey parameters are kept fixed except for the survey area. The results are shown in the left hand panels of Fig. 4.1. Unsurprisingly it is always better to survey a larger area. The lensing uncertainties are inversely proportional to the square root of the sky area and therefore using lensing alone we recover the usual result that the figure of merit (FoM) is directly proportional to the survey area (not shown but see e.g. Amara & Refregier 2007). This is because the figures of merit combine the uncertainties on two measured parameters. Combining with CMB constraints weakens the dependence of FoM on survey area below a direct proportionality (left panels of Fig. 4.1) because the CMB contribution to the constraint is unaffected by the survey area. The exact shape depends on the details of the parameters considered and the observables included.

The black lines show the dark energy figure of merit in General Relativity. As discussed in Chapter 3, we propose that the main results from a cosmic shear survey should combine the information from galaxy shears and galaxy positions, shown by the solid lines throughout this chapter. For reference we also show the conventional approach in which shear correlations alone are used,

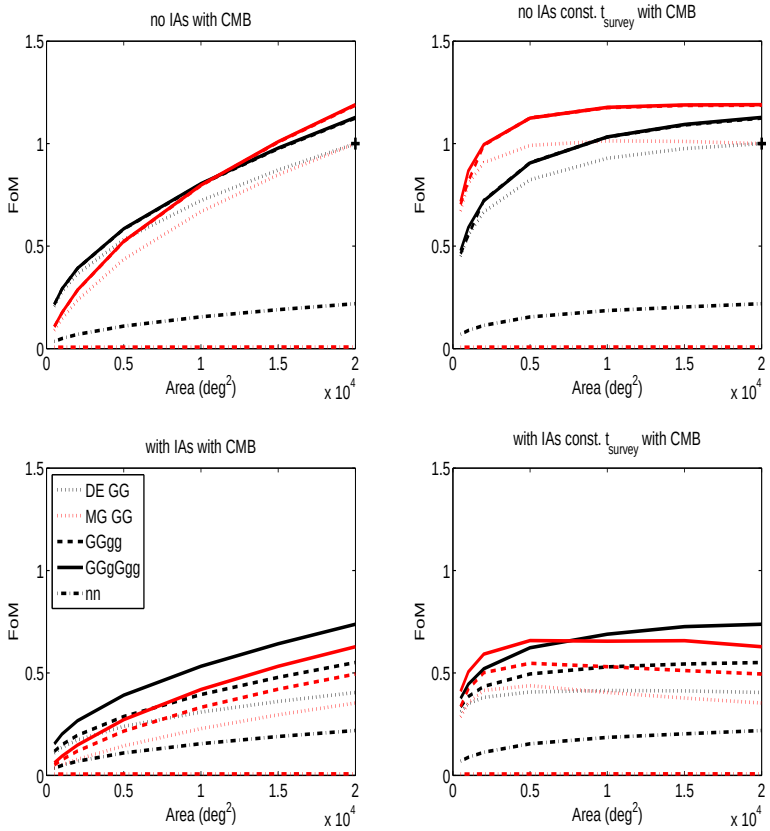

Figure 4.1: Figures showing the variation with survey area of a relative Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [black lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [red lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. The figures show relative FoMs for an optimistic scenario, in which uncertainities in the IA model have been excluded, [top panels] and a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model [lower panels]. The baseline model (relative FoM=1) is shown as a black cross in the top panels. We consider the impact on the relative FoM of increasing survey area, by increasing survey time for fixed limiting magnitude [left panels] and fixing the survey time to trade-off survey area and depth [right panels]. Four data combinations are considered in each panel, shear-shear correlations '$\epsilon\epsilon$' [dotted lines], galaxy-galaxy positions 'nn' alone [dot-dash], '$\epsilon\epsilon$+nn' [dashed] and when shear-position cross-correlations '$n\epsilon$' are included [full lines]. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for these figures in the text.

as dotted lines throughout. In addition we show two other combinations of data: galaxy positions alone (dot-dashed lines) and shear correlations plus galaxy correlations missing the shear-position cross-correlation information (dashed lines). Throughout, we normalise the dark energy figure of merit by dividing by the figure of merit for the conventional set-up for our fiducial survey. Where this survey design exists on a figure we have marked it with a black cross.

For our proposed dataset the dark energy figure of merit is nearly proportional to the square root of the survey area. Extra area helps slightly more than it does when the conventional shear correlations alone are considered. Shear-position correlations make essentially no difference when intrinsic alignments are not included, as discussed in chapter 3. Galaxy position correlations alone are also roughly proportional to the square root of the survey area as opposed to directly proportional, due to the inclusion of the CMB.

The modified gravity figure of merit is shown by the red lines. This is slightly more sensitive to survey area than the dark energy figure of merit when intrinsic alignments are not considered, irrespective of the choice of observable. Galaxy clustering alone is insensitive to modified gravity parameters due to the way we have marginalised over galaxy bias and hence the red dot-dashed lines are constant at zero.

We now consider the case where intrinsic alignment contributions are switched on. In all cases, with and without intrinsic alignments, we have marginalised over free parameters in the intrinsic alignment and bias models with $N_z = N_k = 5$. With the intrinsic alignment parameters the figure of merit are all lower, but the dependence is still close to square root of the area for all methods. The modified gravity figure of merit is still slightly more sensitive to survey area than the dark energy figure of merit.

In reality most surveys have a fixed observing time. This inevitably leads to a trade off between total survey area and survey depth, which impacts both number density of lensed galaxies, n_g, and median redshift achieved, z_m. We use a fitting formula for the number of galaxies per unit redshift per square arcminute

$$n(z) = \Sigma_0 \times \frac{3z^2}{2z_0^3} \exp\left[-\left(\frac{z}{z_0}\right)^{3/2}\right] \tag{4.16}$$

(e.g. Baugh & Efstathiou 1993) and we interpolate the numbers from Table 1 of Blake & Bridle (2005) using

$$z_0 = 0.055(r_{\rm lim} - 24) + 0.39 \tag{4.17}$$

$$\Sigma_0 = \frac{35400}{60^2}\left(\frac{r_{lim}}{24}\right)^{19}. \tag{4.18}$$

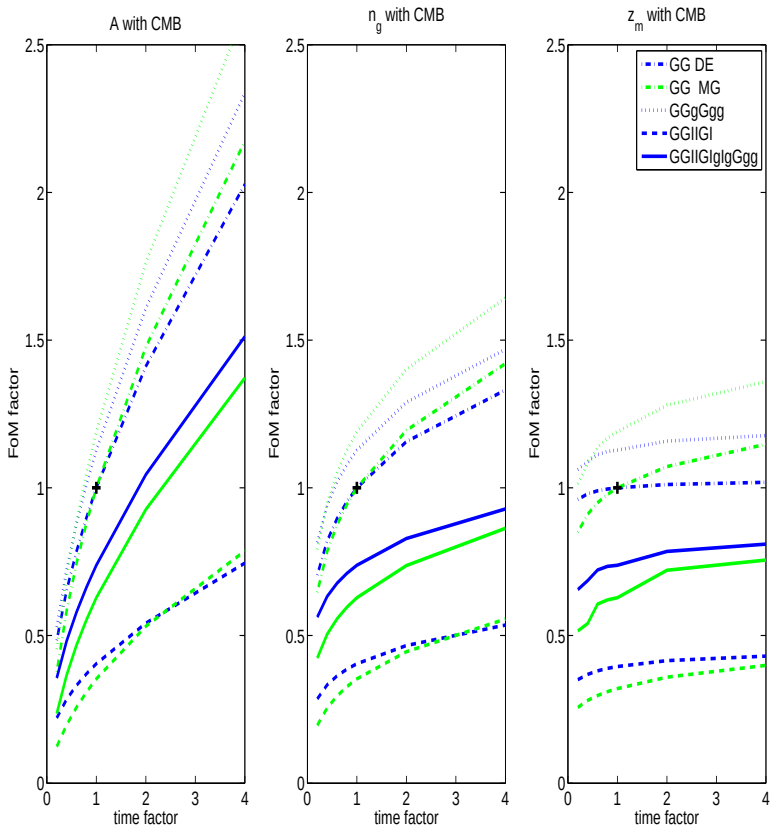

Figure 4.2: Figures showing the variation with time factor of a relative Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [blue lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [green lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. Each figure shows relative FoMs for two scenarios: the first optimistic, in which uncertainities in the IA model have been excluded, for both shear-shear correlations '$\epsilon\epsilon$' alone [dot-dashed lines] and the full combination of shear and galaxy position information '$\epsilon\epsilon$+$n\epsilon$+nn'[dotted lines]. The second, a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model, again for '$\epsilon\epsilon$' [dashed lines] and '$\epsilon\epsilon$+$n\epsilon$+nn' [solid lines]. The baseline model (relative FoM=1) is shown as a black cross. The t-factor represents some change in the amount of available survey time relative the the fiducial stage-IV survey, e.g. t-factor=2 means double the time is available. Results are shown for cases where this extra time is devoted entiely to increasing survey area [left panel], galaxy density [centre panel] and median survey redshift [right panel]. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for these figures in the text.

We use the above equation to scale the number density of galaxies with area relative to the fiducial values. We vary limiting magnitude with area in the usual way as

$$\Delta_{\rm mag} = -2.5 log_{10}(\frac{A}{A_{\rm fid}}) \tag{4.19}$$

and use this to read off the appropriate Σ_0 from Table 1 of Blake & Bridle (2005) which can then be converted into an n_g value. As such we decrease n_g as survey area increases from the fiducial value and vice-versa. As such we are reducing the number density of galaxies in all tomographic bins by the same proportion leaving the distribution of galaxies across redshifts intact. A more detailed future analysis could implement a more sophisticated version of this process where a decrease in depth with increased area would begin by excluding the least bright galaxies which are generally at higher redshift.

Results for varying area with finite survey time are shown on the right hand panels of Fig. 4.1. As expected, the benefits of going to larger survey area are now less strong because of the lost depth. The standard result for cosmic shear analyses is that it is still better to go to larger survey area despite the lost depth. We see that although this is still true at very small survey area, it is not always the case as the survey area increases towards the whole extra-galactic sky. Amara & Refregier 2007 Figure 4 suggests the FoM is roughly proportional to the square root of the area, for fixed observing time without the CMB. Our conventional analysis without intrinsic alignments and using shear correlations alone is shown by the dotted lines in the top right panel of Fig. 4.1). The dependence is now significantly flatter than a half power, which we find is due to the addition of the CMB. We also see that while the dark energy figure of merit does keep increasing to larger survey areas, if weakly, this is not the case for the modified gravity figure of merit, which plateaus beyond about 5000 square degrees. When the galaxy position information is added the same story holds, but the dependence on area is slightly stronger (solid lines in the top right panel of Fig. 4.1).

When intrinsic alignments are added (bottom right panel of Fig. 4.1) the plateau effect is stronger still. Even when just using the shear correlations alone (dotted lines), the dark energy and modified gravity figures of merit are quite flat with survey area, from a few thousand square degrees onwards. The modified gravity figure of merit even peaks slightly at 5000 square degrees. This is because intrinsic alignments dominate over cosmic shear at very low redshift, and by marginalising out the intrinsic alignment nuisance parameters we effectively remove the information at low redshift. Therefore we expect any cosmological figure of merit to eventually reduce in the (unphysical) limit of very large area and small depth.

The galaxy clustering alone constraints (dot-dashed lines) are insensitive to the intrinsic align-

ments and therefore increase as usual. When using the full dataset there is some benefit to the dark energy figure of merit (solid black line) of going to the maximum survey area of 20000 square degrees, although a 5000 square degree survey gathers 80% of the information. The modified gravity figure of merit (solid red line) again peaks around 5000 square degrees and drops by a few percent to the maximum area. Note that the exact details of these conclusions will depend on the fiducial intrinsic alignment model and the number of intrinsic alignment and bias nuisance parameters marginalised over. The results using just the shear-shear and galaxy position- galaxy position correlations (ignoring the shear-position cross term) are shown by the dashed lines and are lower than the full dataset but bear out the same conclusions.

We try to disentangle the contributing factors that vary when the total survey time is kept fixed, following the method in Amara & Refregier 2007. We first consider the case when increasing the survey time is used to increase the survey area, shown in the left hand panel of Fig. 4.2. Next we use the scaling relation between survey time and number density of galaxies $n_g \propto t_s^{0.44}$ (Amara & Refregier 2007) to isolate the effect of changing n_g alone, shown in the center panel of Fig. 4.2. Finally we isolate the survey median redshift, which changes with survey time approximately according to $z_m \propto t_s^{0.067}$ (right hand panel of Fig. 4.2).

We can see that for all scenarios considered (shear-shear without IA through to all data with IA) the dominant effect is the change in survey area, followed at some distance by the number density of galaxies. Changing the survey parameters generally has the biggest effect on the modified gravity normalised figure of merit, as seen in the left hand panels of Fig. 4.1.

4.4 Photometric Redshifts

Accurate redshift information is essential for weak lensing tomography. As survey size increases the number of redshifts becomes of order hundreds of millions to billions. The expense of acquiring spectroscopic redshifts for this number of galaxies is prohibitive. As a result the next generations of WGL surveys will rely on photometric redshifts.

In this chapter we follow the photometric redshift model of Amara & Refregier (2006), including the statistical dispersion of measured photometric redshifts and catastrophic failures in redshift measurement. The statistical dispersion is modelled assuming that measured photometric redshifts, z_{phot} are described by a Gaussian probability distribtion centred on the true redshift, z_t. The width of the Gaussian is the photometric redshift error which we assume scales as $(1+z)$ and

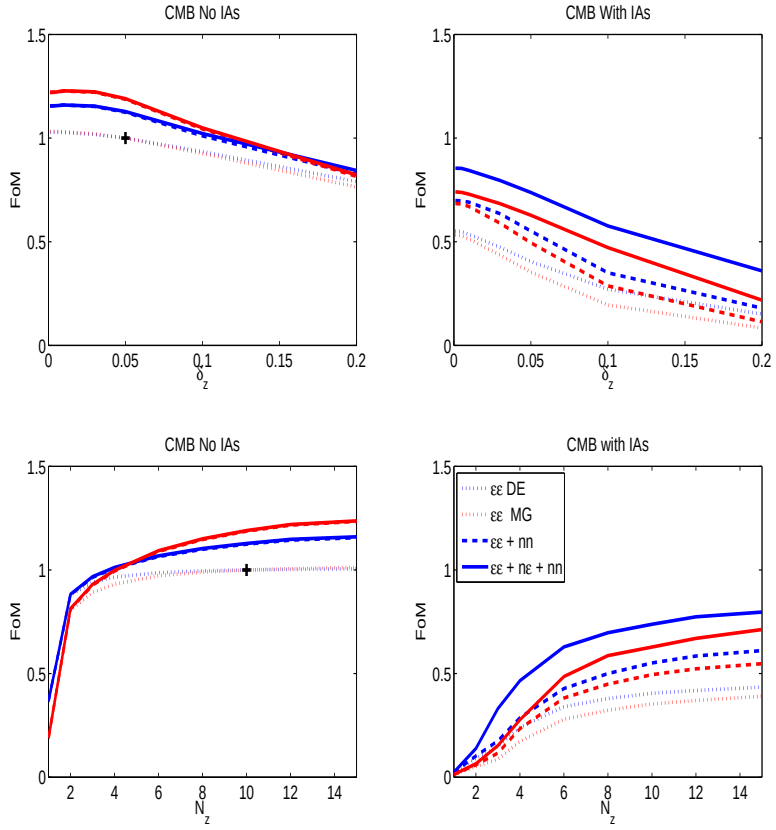

Figure 4.3: Figures showing the variation with photometric redshift error (δ_z) [top panels] and number of tomographic bins (N_z) [bottom panels] of a Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [blue lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [red lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. The figures show relative FoMs for an optimistic scenario, in which uncertainities in the IA model have been excluded, [left panels] and a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model [right panels]. The baseline model (relative FoM=1) is shown as a black cross in the left panels. Three data combinations are considered in each panel, shear-shear correlations '$\epsilon\epsilon$' [dotted lines], '$\epsilon\epsilon$+nn' [dashed] and when shear-position cross-correlations '$n\epsilon$' are included [full lines]. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for these figures in the text.

is controlled by a paramter δ_z,

$$P_{z_{phot}}(z) = \frac{1}{\sqrt{(2\pi)\delta_z(1+z)}} \tag{4.20}$$

In addition to this uncertainty in the photometric PDF galaxy redshifts can be entirely misidentified if, for example, the wavelength range examined is insufficient to identify important features in the spectrum. A misidentified redshift can be said to be assigned a redshift offset by an amount Δ_z from the true redshift. Affected galaxies are known as catastrophic outliers.

We construct a PDF to describe these catastrophic outliers, $P_{z_{cat}}$, and combine it with the scatter PDF, writing the full probability distribution

$$P(z_{phot}|z_t) = (1 - f_{cat})P_{z_{phot}}(z_t) + (f_{cat})P_{z_{cat}}(z_t) \tag{4.21}$$

where f_{cat} is the fraction of our measured galaxies which suffer catastrophic redshift estimation failures.

The distribution of $P_{z_{cat}}$ is bimodal, reflecting the fact that galaxies can either be misidentified as being higher or lower than their true redshifts,

$$< P_{z^{\pm}_{cat}} >= z^{\pm}_{cat} = z_t \pm \Delta_z \tag{4.22}$$

where $P_{z_{cat}} = P_{z^{-}_{cat}} + P_{z^{+}_{cat}}$. The value of Δ_z will depend on the survey filters, photometric reconstruction technique and spectral properites of the survey galaxies. We assume a fiducial value of $\Delta_z = 1$ and that the uncertainty on the PDF of the population of catastrophic outliers is the same as for the slightly scattered galaxies

$$\sigma(P_{z^{\pm}_{cat}}) = \delta_z(1 + z^{\pm}_{cat}). \tag{4.23}$$

This scatter propagates through our Fisher Matrices for the purposes of forecasting by broadening the weight functions for each tomographic bins. Broader bins allow less scope for constraining redshift variation in our various observables, weakening forecast constraints.

Fig. 4.3 shows the variation in FoMs with δ_z and N_z, the number of tomographic redshift bins, for different combinations of observables, with and without IAs. There is a parallel between varying δ_z and N_z- lower redshift error focuses the redshift distribution per bin into a tighter shape, allowing us to learn from changes with redshift over shorter intervals. Somewhat equivalently increasing the number of tomographic bins gives us, in principle, the opportunity to detect changes with redshift at higher resolution.

Increasing δ_z decreases the constraining power of a survey because the loss of redshift precision "smears out" the redshift distribution of our sample. In particular galaxies leak between

tomographic redshift bins meaning we can extract less information from tomography. The top left hand panel shows the effect on various probe combinations, without including the effects of intrinsic alignments. The effect is relatively minor, especially for $\epsilon\epsilon$ alone (dotted lines) whose DE FoM drops by only 1/4 from $\delta_z = 0$ (spectroscopic, i.e. perfect redshift precision) to $\delta_z = 0.2$. This is a much higher value than the fiducial redshift uncertainties of typical Stage-III (0.07) and Stage-IV (0.05) experiments. For each of the probe combinations, DE and MG FoMs display very similar behaviour. The trends become slightly steeper when galaxy position information is included (dashed and solid lines).

The reasons for this behaviour are relatively well understood. The lensing integral (eqn. 4.4) modulates the 3D matter power spectrum by the lensing weight functions $W_i(\chi)$ of the redshift bins being correlated. These lensing weight functions are broad in redshift and act as kernels smoothing the redshift information. This effect limits the usefulness of increasingly accurate redshift information and explains why the trend with δ_z is relatively shallow. Galaxy position data enters the angular power spectrum integral modulated by the redshift distribution $n_i(\chi)$ of the bin in question. This is a less broad function, providing more scope for improvement in accuracy with better redshift information, explaining the steeper lines.

The slightly steeper curves for the MG FoMs may come from the larger number of parameters that need to be constrained in this model, which can be done with new types of information such as increased redshift resolution. Broadly the results for modified gravity are very similar to those for dark energy.

The inclusion of IA terms for the same probes in the top right panel shows the expected decrease in constraining power as we have to marginalise over IA bias terms. As we have seen previously, the MG FoMs take more of a hit due to the IA effects than their DE counterparts.

The decrease in FoM with increasing δ_z is significantly more pronounced once IAs are included, with $\epsilon\epsilon$ including IAs losing $\sim 67\%$ of its constraining power for DE over the interval probed. Qualitatively we might hope to understand this from the fact that the cosmic shear and each IA term, II & GI, have a different redshift dependence. Improved redshift knowledge allows us to better discriminate between the cosmic shear signal and IA contamination. This produces stronger constraints on the IA nuisance parameters and consequently tighter constraints on DE and MG.

All of the curves with varying number of tomographic redshift bins, N_z, show the characteristic plateaux behaviour (bottom left panel of fig. 4.3). Without IAs, increasing the number of tomographic bins becomes much less effective after $N_z \sim 2-3$. After this number $\epsilon\epsilon$ is effectively

flat, while $\epsilon\epsilon + nn$ and $\epsilon\epsilon + n\epsilon + nn$ flatten out slowly, gaining little benefit after $N_z \sim 10 - 12$. This behaviour is linked to the physics described above- as the lensing integral, and in this case photo-z scatter, smears-out redshift information there is a limit to how useful it is breaking up our survey range into ever smaller slices in redshift.

The presence of IAs (bottom right panel) makes the plateaux behaviour less pronounced and pushes the levelling off of FoM back towards $N_z \sim 6 - 10$ for all probes. The use of smaller slices in redshift space gains information about the z-distribution of the IA signals and aids our constraining power. The effect is curtailed by the same fundamental limit of redshift accuracy in the projected angular power spectrum. Again the hit on FoM due to IAs is more pronounced for MG than DE but the qualitative behaviour of FoM with N_z in the presence of IAs is the same for both.

The effect of changing the fraction of catastrophic outlier redshifts in our survey sample is shown in Fig. 4.4. As we would expect increasing the fraction of completely mis-estimated redshifts makes for poorer constraints on both MG and DE, for similar reasons to the impact of δ_z.

All the lines in Fig. 4.4 are relatively flat, suggesting that the projected nature of the angular power spectra and the effect of the lensing weight function continue to limit the impact of redshift information on overall constraints. No probe combination loses more than 30% of its DE or MG FoM as f_{cat} increases from zero to 0.2 which is more than twice the typical requirement value of projected stage-IV surveys.

The general pattern that MG FoMs decrease slightly more strongly than their DE equivalents with the addition of IAs (right panel) is repeated here.

One interesting feature is that the MG constraints for $\epsilon\epsilon$ without IAs and all the probe combinations with IAs decrease more steeply with increasing f_{cat} than their DE counterparts. The difference is most pronounced for $\epsilon\epsilon$ with IAs where the MG constraints falls by twice as much over the f_{cat} range as the DE FoM. It is strongly possible that the extra redshift information helps to constrain the MG parameters because of the specific redshift dependence of the MG trigger parameters.

4.4.1 Redshift Priors

As well as describing the change in FoMs with the size of the photometric uncertainty, it is important to the quantify the effect of our *degree of knowledge* of that uncertainty on our FoMs. In our case we follow the standard procedure of inserting this degree of knowledge as a prior on the parameter δ_z.

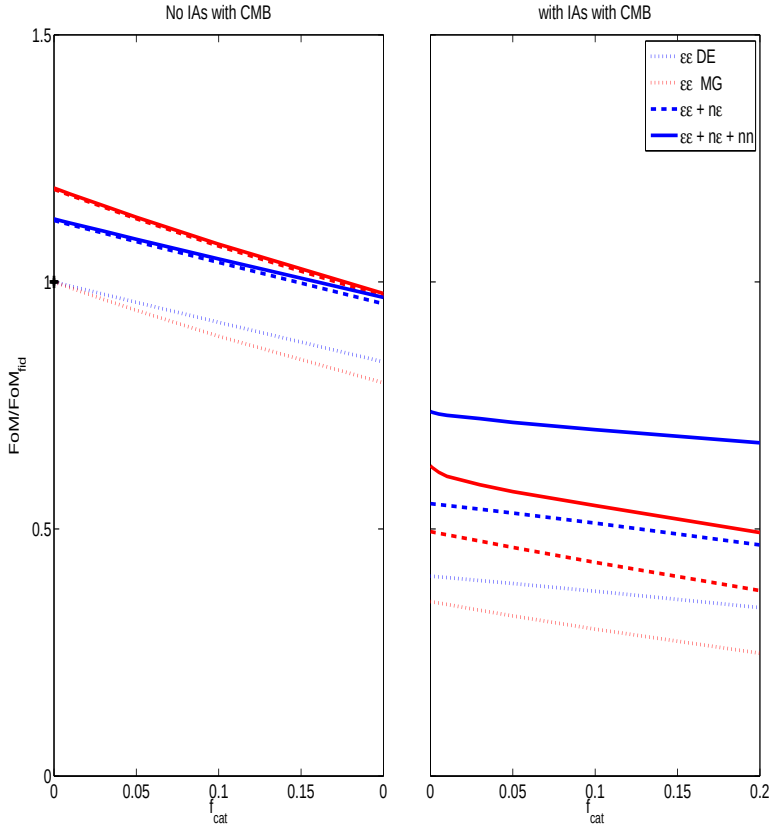

Figure 4.4: Figures showing the variation with fraction of catastrophic redshift outliers (f_{cat}) of a Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [blue lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [red lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. The figures show relative FoMs for an optimistic scenario, in which uncertainities in the IA model have been excluded, [left panel] and a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model [right panel]. The baseline model (relative FoM=1) is shown as a black cross in the left panel. Three data combinations are considered in each panel, shear-shear correlations '$\epsilon\epsilon$' [dotted lines], '$\epsilon\epsilon$+nn' [dashed] and when shear-position cross-correlations '$n\epsilon$' are included [full lines]. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for these figures in the text.

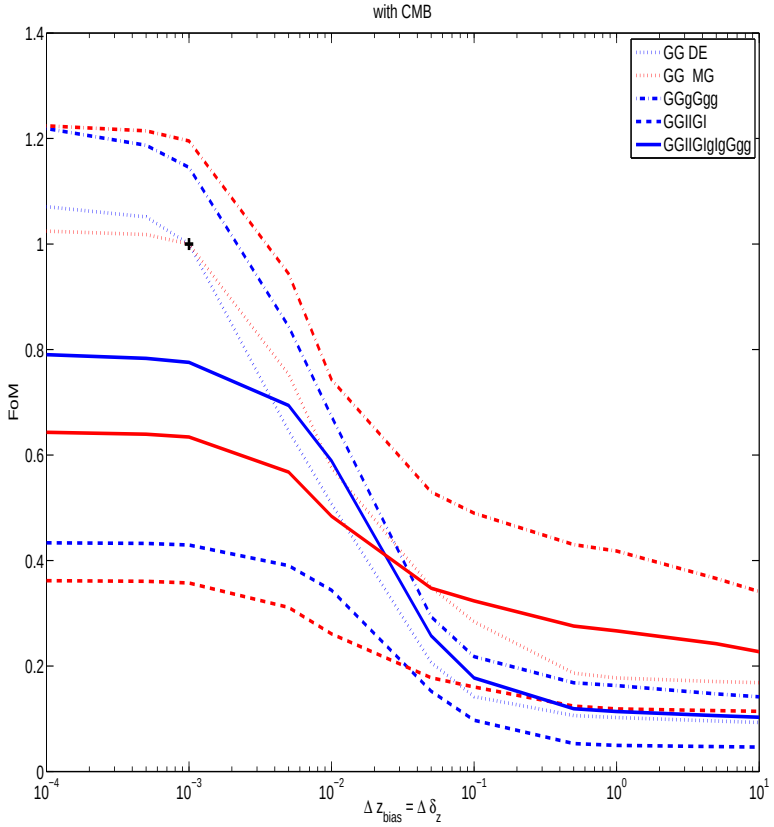

Figure 4.5: Figure showing the variation with the priors on the bias on median redshift, z_{bias}, of 30 redshift bins spanning the survey redshift range due to the effect of photometric redshift uncertainties and the photometric redshift statistical dispersion, δ_z. All other survey parameters are fixed at their fiducial values. Results are for a Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [blue lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [red lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. Two data combinations are considered, shear-shear correlations '$\epsilon\epsilon$' for both an optimistic scenario, in which uncertainities in the IA model have been excluded, [dotted lines] and a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model [dashed lines] and the full combination of shear-shear and galaxy position information including cross-correlations '$\epsilon\epsilon$+$n\epsilon$+nn', again ignoring [dot-dashed lines] and including [solid lines] IAs. The baseline model (relative FoM=1) is shown as a black cross in the left panel. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for this figure in the text.

To test this assumption we extend our photometric redshift model- breaking the redshift range 0-3 into 30 bins in z of with 0.1, following Ma, Hu & Huterer (2005). Note that these bins are used only to parameterise our photo-z model, we continue to use the standard 10 bins for shear tomography. We allow δ_z to vary independently in each of these z-bins and introduce a new parameter $z_{\rm bias}$, the amount by which the the photometric distribution of redshifts is offset from the true redshift. $z_{\rm bias}$ is allowed to vary independently in each of our 30 z-bins. Each δ_z has the standard fiducial value of 0.05 and each $z_{\rm bias}$ has fiducial value zero. We have enlarged our Fisher Matrix by 60 new free parameters, all of which are marginalised over in the results given in Fig. 4.5.

We assume that the Gaussian prior assigned to each of the 60 free parameters is the same and we allow the prior to vary from a width of 0.0001 to 10. The results in Fig. 4.5 show that the constraining power a given survey achieves decreases as we reduce our amount of prior knowledge about the redshift distribution.

The results show a number of interesting effects which vary significantly with different probe combinations and between DE and MG FoMs. Before discussing individual results in detail it is worthwhile noting some common properties and trends across the probes.

Apart from the expected decrease with wider priors we see two distinct regimes- relatively flat FoMs below prior ~ 0.001 and above 1 with a transition regime in between. It is reassuring to note that we cannot achieve arbitrarily large FoMs with tighter redshift priors- this is connected to the previously noted "smearing" of redshift information in the shear angular power spectrum. At large values of the prior width we reach a "self-calibration" regime in which the information in the survey is sufficient to constrain the unknown parameters (Huterer *et al.* 2006).

The results without IAs (dotted and dot-dashed lines) undergo a bigger drop than the corresponding results with IAs (dashed and solid lines). Mostly the lines start higher at narrow prior width and converge to a similar low FoM at large prior width. The lines without IAs are higher at narrow prior width because there are fewer nuisance parameters marginalised over, and so they have further to fall.

What is also clear is that the DE FoM falls off much more steeply over the prior range than the MG FoM. This is true for all probe combinations, with and without IAs but is most spectacular for the lines which include IAs, here there is actually a cross-over in the solid lines, with the MG FoM ending higher than the DE FoM at prior = 10. Focussing on the lines that include IAs, the cosmic shear only lines (dashed) show a factor of 10 decrease in DE FoM but a factor of 3 decrease in MG FoM. This trend is repeated when all data is used (solid). We expect the dark energy parameter

w_a to be particularly sensitive to incomplete redshift information because it describes the time evolution of the dark energy equation of state. By contrast, the modified gravity parameters we use are constant with redshift and the redshift evolution of the modified gravity functions is fixed by our choice of redshift power law s.

4.5 Spectroscopic redshifts

Our discussion of redshift quality has thus far assumed that all our redshift information comes from photometric redshifts, gathered as part of an optical cosmic shear survey. This is certainly the most likely source of redshift information for the large numbers of source galaxies which will be probed by stage-III and stage-IV surveys. Photometry is relatively inexpensive and the information comes from the same images used to measure galaxy shapes.

However, there will be at least some spectroscopic redshift information used in future lensing surveys. It is much more accurate than photo-z (we can approximate $\delta_z^{\rm spec} = 0.0001$) but also a great deal more expensive and time consuming. It also requires a spectroscope instrument to take measurements which must be matched to optically identified galaxies. Nevertheless, any survey relying on photo-zs will still use a small spec-z "calibration sample", generally representative of the overall survey.

Beyond this, it is an interesting question to consider how including some spec-z information in our calculations could affect our constraints. There are a number of dedicated spec-z surveys being planned or carried out at present, see chapter 5 for more details. Cost and time constraints mean they are necessarily of smaller area than our fiducial stage-IV survey.

In practical terms the spec-z quality of redshifts is introduced to our standard FM formalism by setting the photo-z scatter to an extremely low level ($\delta_z = 0.0001$) and the fraction of catastrophic outliers, f_{cat} is set to zero. This allows a perfectly sharp redshift distribution (within the limits of a finite number of tomographic bins) and generally improves the quality of constraints.

The upper panel of Fig. 4.6 shows DE and MG FoMs for different probe combinations as we add spectroscopic survey area to our fiducial photo-z survey. A_{spec} denotes the amount of purely spectroscopic survey area included in our calculations. All lines are normalised to the fiducial photo-z survey for that particular probe/IA combination, i.e. the point $A_{spec} = 0$ on the x-axis.

As expected, adding extra survey area improves the FoM. The trend is rather shallow though, with $\epsilon\epsilon$ and $\epsilon\epsilon + n\epsilon + nn$ without IAs improving only by 5% at $A_{spec} = 2000 deg^2$ even though the combined area of the survey has increased by 10%. Probes including IAs show much stronger

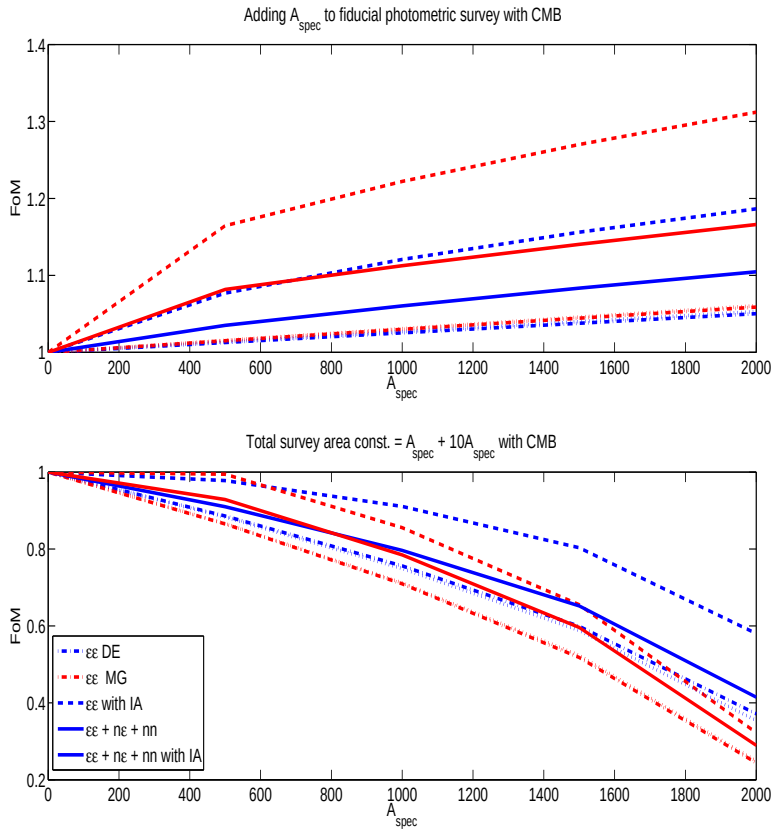

Figure 4.6: Figure showing the variation with the amount of spectroscopic area, A_{spec}. Results are for a Stage IV survey FoM for dark energy (DE) parameters w_0, w_a [blue lines], and modified gravity (MG) parameters $Q_0, Q_0(1+R_0)/2$ [red lines], as described in Eqn. 4.2 and Eqn. 4.3, relative to a 'baseline' FoM. The baseline includes shear auto-correlations 'GG' alone, excluding IAs, over a 20,000 square degree survey. All FoMs contain priors from a Planck-like CMB survey. Two data combinations are considered, shear-shear correlations '$\epsilon\epsilon$' for both an optimistic scenario, in which uncertainities in the IA model have been excluded, [dotted lines] and a conservative scenario, in which uncertainties in the IA model are marginalized over using a $n_k = n_z = 5$ gridded bias model [dashed lines] and the full combination of shear-shear and galaxy position information including cross-correlations '$\epsilon\epsilon$+$n\epsilon$+nn', again ignoring [dot-dashed lines] and including [solid lines] IAs. The baseline model (relative FoM=1) is the $A_{spec} = 0$ point on each panel. Two scenarios are considered, the first adds pure spectroscopic area to the fiducial survey [top panel] while the second employs a tradeoff where photometric survey area decreases as spectroscopic area is added at a rate of 10deg^2 of photometric area for every 1deg^2 of spectroscopic area [bottom panel]. For the galaxy position correlations a $n_k = n_z = 5$ gridded bias model is used throughout. We discuss the conclusions for this figure in the text.

improvements- particularly in MG FoM, with $\epsilon\epsilon$ including IAs improving by 30% over the A_{spec} range. This is consistent with our previous findings- there is only so much improved redshift information can do to benefit the lensing integral. IAs and MG allow more room for improvement as better redshift knowledge helps to discriminate with different evolutions of the power spectra and trigger parameters.

Note that Fig. 4.1 showed us that FoM improved roughly as the square root of the area. For a 10% increase in area, even if the extra was still photometric, we would expect an increase of a factor of $\sim$3.2 in FoM. We see this for the non-IA curves, suggesting that they are not benefitting from the spectroscopic nature of the new data, simply from the extra area. On the other hand, the probes including IAs do seem to get an extra benefit from the spectroscopy, especially for MG, in particular for $\epsilon\epsilon$ with IAs.

It's clear that adding area, spectroscopic or photometric to a given survey will help improve constraining power. A slightly different approach asks what are the effects if we trade photometric coverage for more expensive spectroscopic area. We try to begin answering this question in the lower panel of Fig. 4.6. Assuming a fixed total observing time equivalent to our fiducial photo-z survey of 20,000deg^2 we make the approximation that spectroscopic are takes ten times as long to cover as the same photometric area. As we increase the proportion of spectroscopic coverage, $A_{\rm spec}$, we vary the photometric area according to 20,000deg^2 $= A_{\rm phot} + 10A_{\rm spec}$. For example, points at $A_{\rm spec} = 1000$deg^2 represent joint constraints from 1,000deg^2 of spec-z area and 10,000deg^2 of photo-z area.

No probe combination benefits from including spec-z area in the mix at this cost level, either for DE or MG FoMs. Probe combinations including IAs benefit most from the increased use of spec-z area, with $\epsilon\epsilon$ inc. IAs FoMs remaining almost flat until A_{spec} 500deg^2 and dropping thereafter. Other probe combinations, without IAs, drop as soon as the total area is reduced.

One interesting feature is that MG FoM drops more steeply than the DE FoM as A_{spec} is increased, particularly for $\epsilon\epsilon$ with IAs. This is in strong contrast with the upper panel of Fig. 4.6 where the MG FoMs benefitted most from the inclusion of spec-z area. The results tell us that the increased redshift accuracy is greatly outweighed by the penalty from the reduction in overall area.

Due to constraints on computational time, our spectroscopic data were analysed with $N_z = 10$, the same number of tomographic bins as the photometric survey data. Preliminary results show that, with spectroscopic redshifts, the FoM continues increasing with N_z after the plateaux seen in photometric results. For a fair comparison we should analyse the spec-z information with $N_z > 10$

to get the most from the spectroscopy, then combine. This is an approach we plan to investigate in future work.

4.6 Conclusions

We have confirmed that the principle results of chapter 3 are robust across a wide range of survey parameters. The inclusion of IA effects, and marginalisation over a realistic model for biasing due to IAs, significantly degrades the constraining power of a cosmic shear survey studying either dark energy or deviations from general relativity. The inclusion of galaxy position information through position-position and position-shear correlations goes some way towards mitigating the effect.

In this chapter we have extended the analysis to consider the impact of some of the most important cosmic shear survey parameters. For each parameter we chart the impact on the DETF FoM and our analogously defined MG FoM.

Whether we ignore or include intrinsic alignments, increasing survey area improves our FoMs for both DE and MG. This is expected but, contrary to previous work which did not include the CMB, we find that including CMB priors weakens the dependence on area to one roughly proportional to the square root of the area, rather than the linear variation seen without the CMB.

When we place a more realistic constraint on our ability to increase survey area we see markedly different behaviour. Assuming our observing time is fixed we include a trade-off between increased area and number density of galaxies achieved in our sample. In this case we see that increasing the area of a cosmic shear survey above a few thousand degrees squared has very little impact on the FoMs achieved for all probe combinations, particularly when intrinsic alignments are included. We can explain this new result by considering that wide shallow surveys pick up a lot of the intrinsic alignment signal because intrinsic alignments dominate at low redshifts. Conversely a deeper narrower survey will gather more information from the higher redshift lensing signal. For MG constraints the FoM even drops at large areas.

We investigate the impact of redshift uncertainty on our ability to measure dark energy and modified gravity. Planned near and medium term cosmic shear missions will rely on photometric redshifts, which are less accurate but cheaper and quicker to collect than redshifts from spectroscopy. The accuracy of our photometric redshifts is parameterised by δ_z, the redshift dispersion parameter. Unsurprisingly as we increase this error our tomographic analysis loses constraining power. However, the decrease in FoM with increased δ_z is relatively shallow due to the smoothing effect of the broad lensing kernel and the projection effects of the 2D angular power spectrum. No

probe combination loses more than a quarter of its constraining power, relative to perfect redshift information, out to $\delta_z = 0.05$, the expected value for stage-IV WGL experiments.

A similar story is told when we vary N_z, the number of tomographic redshift bins used in our analysis. The smearing out of redshift information by the lensing kernel means that FoMs tend to plateau arount $N_z \sim 4-6$, there is only so much to be gained from ever-finer redshift slicing. The addition of IAs weakens this effect somewhat, pushing the plateau back to $N_Z \sim 8-10$. This is expected because the detailed redshift information is used to disentangle the very localised II signal from the specific redshift template for the GI signal and the cosmological lensing signal.

We have presented results showing the effect of varying our degree of prior knowledge of an extended parameterisation of redshift information, allowing 60 free parameters to be shared between redshift dispersion and redshift bias. As we would expect, tight priors produce strong FoMs which weaken as the priors become more loose. Around $\Delta z_{\text{bias}} = \Delta\delta_z = 0.1$ the probes enter a self-calibration regimes and the FoMs flatten off. The DE FoMs fall off much more steeply and relatively lower than their MG counterparts, highlighting the sensitivity of the DE equation of state, particularly w_a, to incomplete redshift information.

There will be a number of dedicated spectroscopic galaxy surveys conducted over the same time-frame as the large photometric surveys we have modelled. We show that the addition of small areas of spectroscopic information to our fiducial photo-z survey does improve FoM. However, with the exception of $\epsilon\epsilon$ alone, including IAs, the improvement is never significantly greater than that which would be produced by the addition of the same extra area with photometric quality redshifts.

When we model a situation where there exists a trade off between the collection of photo-z data and (ten times more costly) spec-z data, we find that no probe combination, for either DE or MG, improves its FoM by adding spectroscopic area to a purely photo-z survey. We have assumed a relatively simple model for deviations from GR. We vary two MG parameters, each with the same fixed redshift dependence. In the future we hope to extend these results to take into account a MG gravity parameterisation which could vary more generally as a function of scale and redshift.

CHAPTER 5

COMBINING IMAGING AND SPECTROSCOPY: THE CASE FOR DES & DESPEC

5.1 Introduction

In the previous chapter we explored the importance of good-quality redshift information for cosmic shear surveys in the presence of IAs and when combined with galaxy position information. We touched briefly on what role fully spectroscopic redshift measurements could play. In this chapter we focus more explicitly on dedicated spectroscopic surveys and some of the science they could produce in combination with broader photo-z surveys.

Spectroscopic redshifts are highly accurate. Light from a source galaxy is split into its frequency components and analysed by a spectrograph. As the light from distant galaxies is redshifted due to the expansion of the Universe, emission and absorption lines in the galaxy spectrum appear shifted in frequency compared to where they would be for a nearby galaxy. Identifying known spectral lines from, for example, abundant elements such as hydrogen allows us to calculate the redshift of the galaxy.

The downside is that spectroscopic redshifts are expensive and time-consuming to collect. Even the most advanced multi-fibre system can only collect hundreds of spectra at a time, as pioneered by 2dF (Lahav *et al.* 2002) and SDSS (Abazajian *et al.* 2009). The sheer number of objects which modern cosmological surveys, such as the stage-III and IV cosmic shear surveys, are designed to collect makes it impossible to spectroscopically identify all redshifts of interest. Instead multi-band photometry is used. This allows spectral templates of galaxies to be fit against the observed bandwidth intensities (Bolzonella *et al.* 2000). Alternatively machine learning methods

such as neural networks can be used if a training set is available, see for example the ANNz software (e.g. Collister & Lahav 2004). There are a number of competing techniques which do this estimation but all produce a redshift for each object and some quantification of the uncertainty of the fit.

This is not to say that no spectroscopic surveys are planned for the future. We have just noted how spectroscopy is required to calibrate photometric redshift pipelines. This can take the form of a dedicated survey, before or after the main survey, to acquire a representative sample of spectrosopic redshifts which well characterises the redshift distribution of the main survey. Alternatively spec-z results from already existing, or already planned dedicated spectroscopic survey catalogues can be used.

Spectroscopic redshift surveys offer powerful cosmological opportunities in their own right. Their expense means they are generally smaller, with a lower number density than equivalent photo-z surveys but the more accurate redshift information allows us to probe physics including Type Ia Supernovae and BAOs to much higher accuracy. The precision that a spectroscopic survey affords also gives us access to new probes of cosmology. Precise measurement in the redshift domain adds Redshift Space Distortions (RSDs) to the measured galaxy power spectrum. These arise due to the peculiar velocities of galaxies. Accessing the galaxy peculiar velocity field via RSDs acts as an additional probe of cosmology. Current photometric redshifts are simply not accurate enough to provide constraints from RSDs. See below for more information. Ongoing surveys include WiggleZ and BOSS. Among the spec-z surveys planned over the next decade are e-BOSS, BigBOSS, HETDEX, DESpec, SuMIRe, VISTA follow-up and SKA.

The Dark Energy Survey (DES) is a 5,000deg^2 photometric redshift survey using the purpose built DEcam camera on the Blanco 4m telescope in Chile. Due to see first light in late 2011, it will combine four probes: WGL, Supernovae, BAOs and Galaxy Clustering to produce some of the strongest constraints on the dark energy figure of merit.

It is in this context that we discuss a proposed spectroscopic redshift survey called DESpec. Envisaged as a spectroscopic follow-up to the photometric Dark Energy Survey mission, it could be used to calibrate the DES photometric redshifts but, more ambitiously, it could be used to improve the science reach of DES and provide complementary information and probes to produce even stronger joint constraints when combined with DES. If the spec-z follow up was carried out within the DES patch of sky there is the possibility that the cross-correlation between the two measurements could further strengthen their joint constraining power. See chapter 6 for details of the DES mission itself.

This chapter forms part of an initial design study for the DESpec proposal, focusing on the benefits a spectroscopic follow up survey would provide for the WGL work of the DES survey, in terms of greater redshift accuracy for cosmic shear measurements or improved constraining power from combining DES WGL with other probes from DESpec. In 5.1.1 we summarise the DESpec concept before discussing the impact of spectroscopic information on the WGL constraints from DES in 5.2. In 5.3 we introduce Redshift Space Distortions as a science probe provided by spectroscopic surveys and discuss joint constraints between WGL and RSDs. In 5.4 we shift the focus from dark energy to modified gravity before concluding in 5.5.

5.1.1 DESpec

The Dark Energy Spectrometer (DESpec) project is in its very early stages. Applications are currently going in to funding agencies. This work forms part of preliminary science studies which will feed into the writing of a white paper on the project. The basic parameters of the design involve a 4000 fibre spectroscope which would be mounted on the same telescope used for DES, the Blanco 4m telescope in Chile. The survey would measure approximately 10 million galaxy spectra using a target list generated from a subset of the DES galaxy catalogue.

A spectroscopic survey on this scale is a very useful science mission in its own right. Excellent redshift information for $\sim$ 10 million galaxies across 5000deg^2 will produce strong estimates of the galaxy power spectrum shape and amplitude in redshift space, allowing for the measurement of radial Baryon Acoustic Oscillations (BAOs) and Redshift Space Distortions (RSDs), discussed in more detail in 5.3 below. The spec-z catalogue produced could provide an excellent calibration set for DES photometric redshifts and improve the performance of DES probes including cluster finding and supernovae. Additional redshift information and complementary probes are excellent science drivers in their own right but the fact that DES and DESpec would be on the same patch of sky opens new doors. The surveys would share overlapping catalogues and cross-correlation effects between physics at different redshifts could provide additional constraining power if correctly modelled.

Clearly the survey strategy of any mission determines the galaxy population which will be measured. At this stage three representative strategies are being considered for DESpec. Each assumes a sky area of 5,000deg^2 and spectroscopic redshift Gaussian errors of $\delta_z = 0.001$ on a total of 10^7 galaxies. The three strategies can be summarised as:

- Sample 1: 100% spectroscopic completeness of the DES galaxies to r=21 mag, with redshift

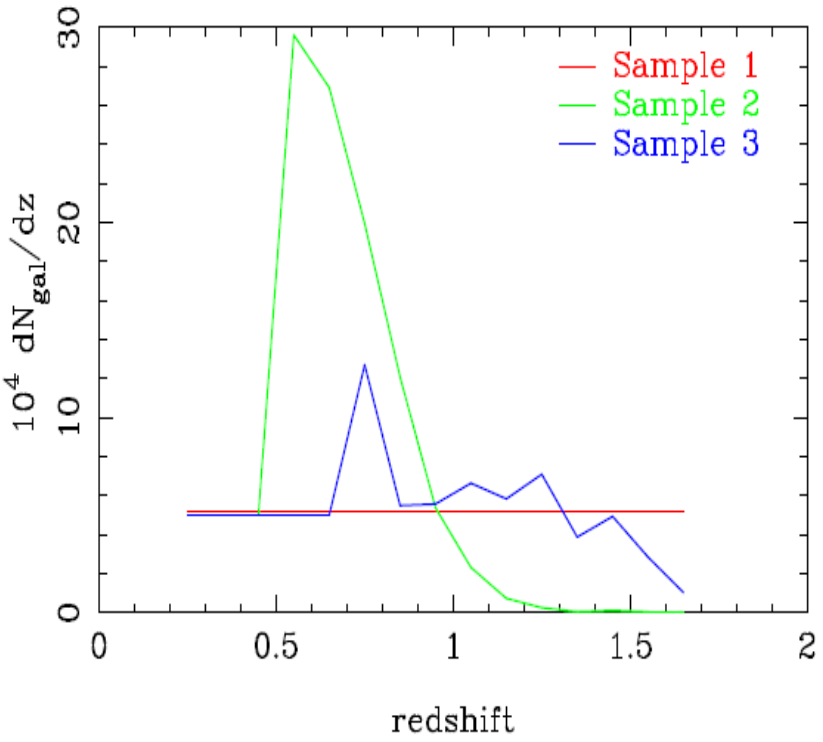

Figure 5.1: The number density of galaxies as a function of redshift in the three samples produced by the proposed survey strategies. (Samushia & Percival, personal communication).

precision of 50 km/sec. Constant number density over the range $0.2 < z < 1.7$.

- Sample 2: The above plus 50% completeness to r=22.5 mag evenly distributed over all redshifts bins above z = 0.5 (using photo-z information to distribute the target selection in redshift). Constant number density $0.2 < z < 0.5$.

- Sample 3: Constant number density $0.2 < z < 0.7$. Constant clustering amplitude, emission line galaxies selected for $0.7 < z < 1.7$, assuming the same S/N cut for redshift determination as the BigBOSS proposal[1]).

The number density of galaxies for each survey strategy are shown in Fig. 5.1. We will assume survey strategy A for the constraints shown in this chapter, comparing with the other strategies in our summary Figure of Merit plots.

Fiducial Survey Paramters	
Parameter	DES
Area	5,000deg^2
n_g	10
σ_γ	0.23
N_z	5
δ_z	0.07
f_{cat}	0
Δ_z	1
z_0	0.8/1.412
α	2
β	1.5

Table 5.1: Fiducial survey parameter values for the DES survey. n_g is the projected number density of galaxies in arcmin^{-2}, σ_γ is the intrinsic ellipticity dispersion of our galaxy sample, N_z is the number of tomographic redshift bins we use to analyse our results. δ_z is the Gaussian photometric redshift error and f_{cat} is the fraction of catastrophically mis-estimated redshifts.z_0, α and β define the galaxy redshift distribution according to 3.34.

5.2 The impact of redshift accuracy on WGL

Weak gravitational lensing is one of the four DES probes of cosmology. By necessity the redshift distribution of galaxies will be determined by photometry. We assume the photometric redshift uncertainty model of Amara & Refregier (2006), as detailed in chapter 4. The photometric redshift dispersion is Gaussian around the true redshift with width $\sigma_z = \delta_z(1+z)$ where $\delta_z = 0.07$. In this chapter we will assume the fraction of catastrophic outliers, $f_{\rm cat}$, is zero. The other fiducial DES survey parameters are given in 5.1, including the parameters which determine the redshift distribution via equation (3.34).

The galaxies targeted by a DESpec-style follow-up survey would be selected from a subset of the DES survey catalogue. As such they effectively provide "upgrades" to some fraction of the DES galaxies with photometric redshifts. We will assume spectroscopic redshifts are measured to an accuracy of $\delta_z = 0.0001(1+z)$ (Weiner *et al.* 2005). It is estimated that DES will survey roughly 280 million galaxies. It is not easy to decide which of these galaxies would have their

[1] `http://cosmology.lbl.gov/BOSS/`

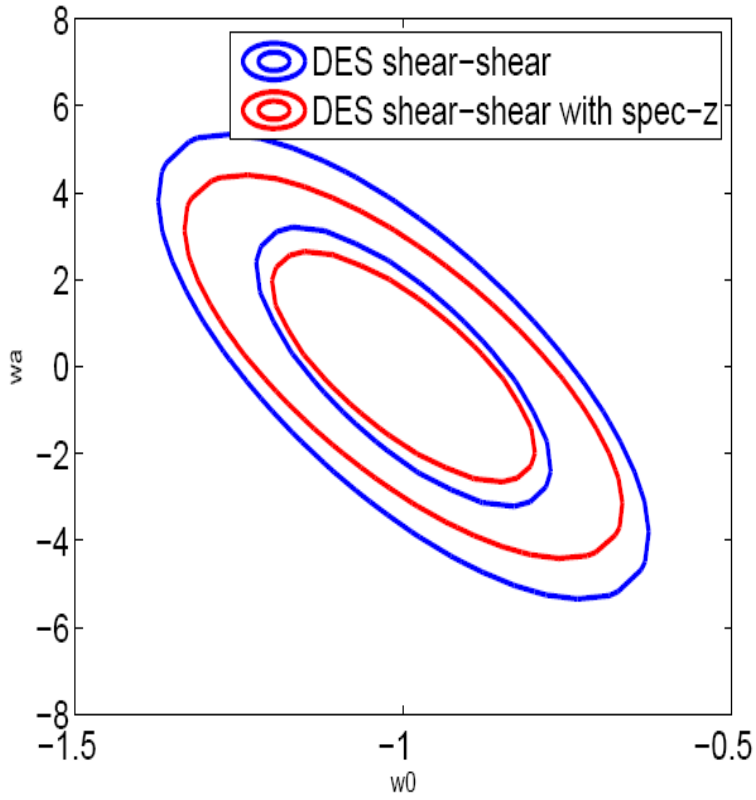

Figure 5.2: Constraints on the dark energy equation of state parameters, w_0 and w_a, marginalised over other cosmological parameters and nuisance parameters. The contours show 68% and 95% confidence limits respectively. Blue contours show the constraints from the standard DES survey set-up with photometrix redshift error $\delta_z = 0.07$. Black contours show constraints for the standard DES survey as if all redshifts were known precisely, $\delta_z = 0.0001$.

photometric redshifts upgraded by being one of the ~ 10 million galaxies recorded by DESpec. We can get some handle on the effect of spectroscopic redshifts on the DES cosmic shear constraints by being wildly optimistic and assuming that all our measured DES galaxies have been assigned spectroscopic-quality redshifts. In practice this means an analysis with no catastrophic outliers, $f_{cat} = 0$, and a very low photo-z scatter, $\delta_z = 0.0001$.

Fig. 5.2 shows the results. We see constraints on the dark energy parameters for DES WGL alone and DES WGL re-analysed with perfect redshift information. These results ignore the effects of intrinsic alignments. The constrains are produced by the Fisher Matrix formalism, varying a set of cosmological parameters: $\{\Omega_m, w_0, w_a, h, \sigma_8, \Omega_b, n_s\}$. The constraints shown marginalise over the other parameters and do not include any of the other DES probes. Wide and flat priors are placed on all parameters and in addition we assume a flat universe.

We see that there is a small improvement in constraining power, $\sim 20\%$ on each dark energy parameter. With a complete DESpec survey we are expecting to get spectra for only $\frac{1}{28}$th of this number of galaxies, suggesting that the direct improvement in constraining power from WGL alone will be very small. This is not unexpected and tallies with the results we saw in the last chapter. The cosmic shear integral

$$C_{ij}(l) = \int \frac{d\chi}{\chi^2} W_i(\chi) W_j(\chi) P_\delta(\frac{l}{\chi}). \tag{5.1}$$

involves the convolution of the matter power spectrum with the broad lensing weight function $W(\chi)$. As we have already noted, this smears out any redshift information and places a limit on the improvement that can be achieved by better redshift accuracy. We have shown that increasing area and pushing photometric surveys deeper are much more effective ways of improving WGL constraints than acquiring spectroscopic redshifts.

5.3 Redshift Space Distortions

A dedicated spec-z survey of 10 million galaxies is a very useful cosmological tool in its own right. Accurate redshift information allows us to reconstruct the galaxy power spectrum to a high quality. Galaxies are biased tracers of the underlying mass distribution so we relate the galaxy power spectrum to the underlying matter power spectrum, subject to some galaxy bias, as $P_g(k,z) = b_g^2(k,z)P_\delta(k,z)$. The amplitude and slope of the matter power spectrum can themselves put some constraints on cosmology. We have mentioned, in the introduction, how the shape of the galaxy power spectrum has a series of "wiggles" imprinted on it from the peaks and troughs of oscillations in the photon-baryon fluid before the epoch of recombination. These Baryon Acoustic Oscillations (BAOs) act as standard rulers, providing a characteristic scale determined by the comoving sound horizon at the drag epoch (shortly after recombination) s. The observed radial BAO scale measures $sH(z)$ in the radial direction and $\frac{D_A(z)}{s}$, the angular diameter distance, in the transverse directions. Measuring the BAOs places even tighter constraints on cosmology (Glazebrook & Blake (2005), Seo & Eisenstein (2003)).

The fact that our survey measures the position of galaxies in redshift space gives us access to a further probe of large-scale structure. Galaxies trace the Hubble flow of the expanding Universe but they also possess a "peculiar velocity" on top of this due to their response to local gravitational potentials. The effect of these peculiar velocities is that the galaxy power spectrum, expected to be isotropic in real space, acquires distortions in redshift space along the line of sight. These Redshift

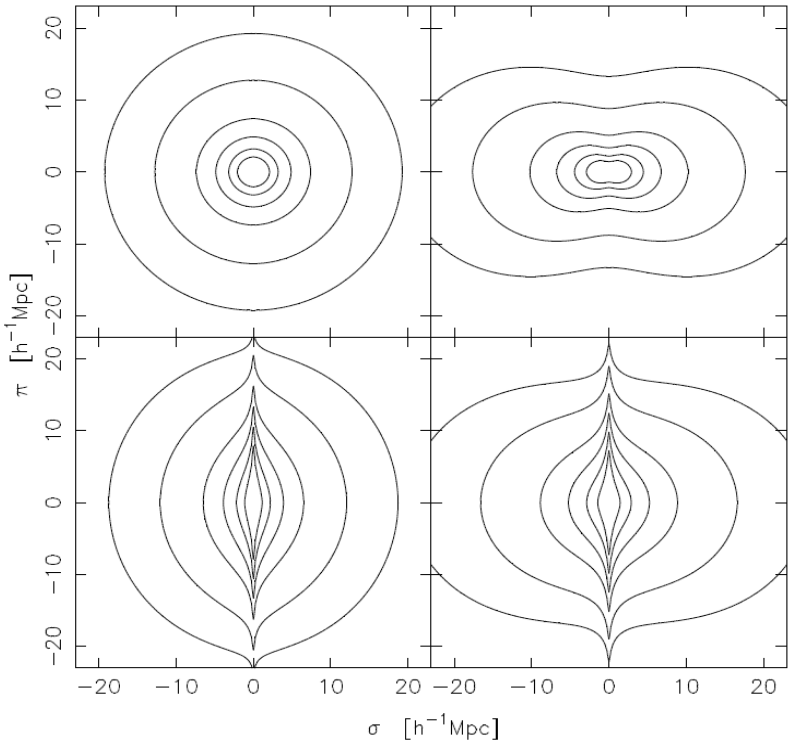

Figure 5.3: Schematic of RSD effects. The top left panel represents an undistorted correlation function, the top right panel is a model with coherent infall added, the bottom left panel is a model with just random pairwise velocities added and the bottom right panel has both infall and random motions added. (Hawkins *et al.* 2003)

Space Distortions (RSDs) cause two major effects: the redshift space galaxy power spectrum, $P_{gg}^{(s)}(k, \mu_k)$ is "squashed" along the line of sight on large scales and elongated on small scales (this is known as the finger-of-God effect) (Kaiser 1988). $\mu_k \equiv k_\parallel/k$ depends on the angle between a wavevector **k** and the line of sight. RSDs are illustrated schematically in Fig. 5.3.

We decompose the redshift space galaxy power spectrum into three components

$$P_{gg}^{(s)}(k, \mu_k) = \left[P_{gg}(k) + 2\mu_k^2 P_{gv}(k) + \mu_k^4 P_{vv}(k)\right] F(k^2\mu_k^2\sigma_v^2) \tag{5.2}$$

(e.g. Guzik *et al.* 2010). We assume a maximum k value for all analysis of $k = 0.15hMpc^{-1}$, keeping us firmly in the linear regime and allowing us to set $F \approx 1$ as F describes non-linear velocity dispersion effects. ν is peculiar velocity field. Note that it is, in principle, possible to treat RSDs using either $P(k, z)$ or tomographic C_l formalisms. For this work we follow the main seam of existing literature and use $P(k, z)$. In the future it may be advantageous to use tomographic C_l^s (which gives us power spectra as a function of redshift, $C_l(z)$, for ease of comparison and

combination with WGL and galaxy position information.

Each of the component auto- and cross-power spectra is obtained via an estimator related the the measures redshift space galaxy power spectrum by a window function:

$$\hat{P}_{XY} = \frac{1}{N_k}\Sigma_{k,\mu}W_{XY}(\mu)\hat{P}^{(s)}_{gg}(k,\mu) \tag{5.3}$$

where $X, Y = \{g, \nu\}$ and

$$W_{gg}(\mu) = P_0(\mu) - \frac{5}{2}P_2(\mu) + \frac{27}{8}P_4(\mu) \tag{5.4}$$

$$W_{g\nu}(\mu) = \frac{15}{4}P_2(\mu) - \frac{135}{8}P_4(\mu) \tag{5.5}$$

$$W_{\nu\nu}(\mu) = \frac{315}{8}P_4(\mu) \tag{5.6}$$

where $P_l(\mu)$ is the l-th order Legendre polynomial. This approach comes from Guzik *et al.* (2010) and is designed to extract an observable for each ndividual power spectrum from a measured redshift space galaxy power spectrum. The summation is carried out over modes contained in spherical shells in Fourier space which satisfy $k_i - \Delta k_i/2 \leq |\mathbf{k}| \leq k_i + \Delta k_i/2$. Our analysis uses 15 logarithmically spaced band powers. We calculate Fisher Matrix constraints using RSDs in the data vector $\hat{D}_\nu = \left\{\hat{P}_{gg}(z_i, k_n), \hat{P}_{g\nu}(z_i, k_n), \hat{P}_{\nu\nu}(z_i, k_n)\right\}$ where the subscript i represents one of five tomographic redshift bins. We assume a constant galaxy bias, $b_g(z)$ within each bin with fiducial value 1.4, each of these acts as a nuisance parameter to be marginalised over. We assume galaxy bias is scale independent. We make the incorrect assumption the three power spectra in our data vector can be combined independently. In a more extended future analysis we will compute the full covariance matrix, including off diagonal terms. With only five bias parameters and zero off-diagonal elements in the covariance matrix this analysis may produce somewhat over-optimistic constraints.

The results of our RSD analysis for a DESpec type survey are shown as constraints on the dark energy parameters in Fig. 5.4. Even when marginalised over cosmology and the galaxy bias parameters, RSD constraints from DESpec are vastly more constraining than DES WGL. Though it should be noted that the DES $\epsilon\epsilon$ constraints shown use broad, flat priors and do not incorporate any information from other probes. There are a number of caveats that we should mention at this stage- the galaxy bias model is not as simplistic as some commonly used but it has a lot less freedom than, for example, our fiducial model from chapters 3 & 4 which employed 26 free nuisance parameters for b_g and allowed it to vary with scale. Future work will attempt to incorporate a looser bias model. We have also made a number of simplifying assumptions, from the independence of the galaxy/velocity power spectra to the implicit assumption of 100%

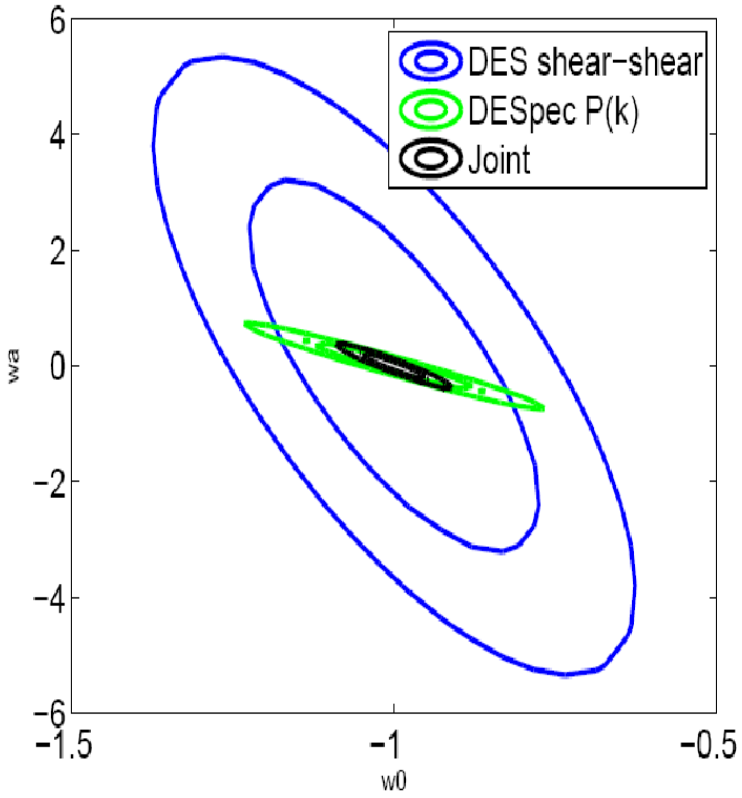

Figure 5.4: Constraints on dark energy parameters, marginalised over other cosmology and galaxy bias nuisance parameters. Inner and outer contours are 68% and 95% confidence limits respectively. Blue contours show constraints for DES WGL alone. Note this analysis has broad, flat priors and does not incorporate any information from other probes. Green contours are DESpec RSDs and red contours are the joint constraints, assuming WGL+RSD can be added independently.

spectroscopic efficiency in our DESpec model- i.e. assuming we actually achieve perfect redshift information for all 10million galaxies. Even so, the powerful nature of RSDs from a dedicated spec-z survey is clear.

Of even greater interest must be the black contour in Fig. 5.4. This is the joint constraint between the DES WGL and the DESpec RSD. It gives a six-fold increase in DETF FoM over the DESpec RSDs alone. The obvious inference is that some parameter degeneracy is being broken by the combination of the probes, leading to a greatly improved joint constraint. Of course, the joint constraint is orders of magnitude stronger than DES WGL alone, confirming the idea that cosmic shear constraints may not benefit from spectroscopic redshifts directly but there exists strong incentives to produce spec-z surveys for the purposes of joint constraints.

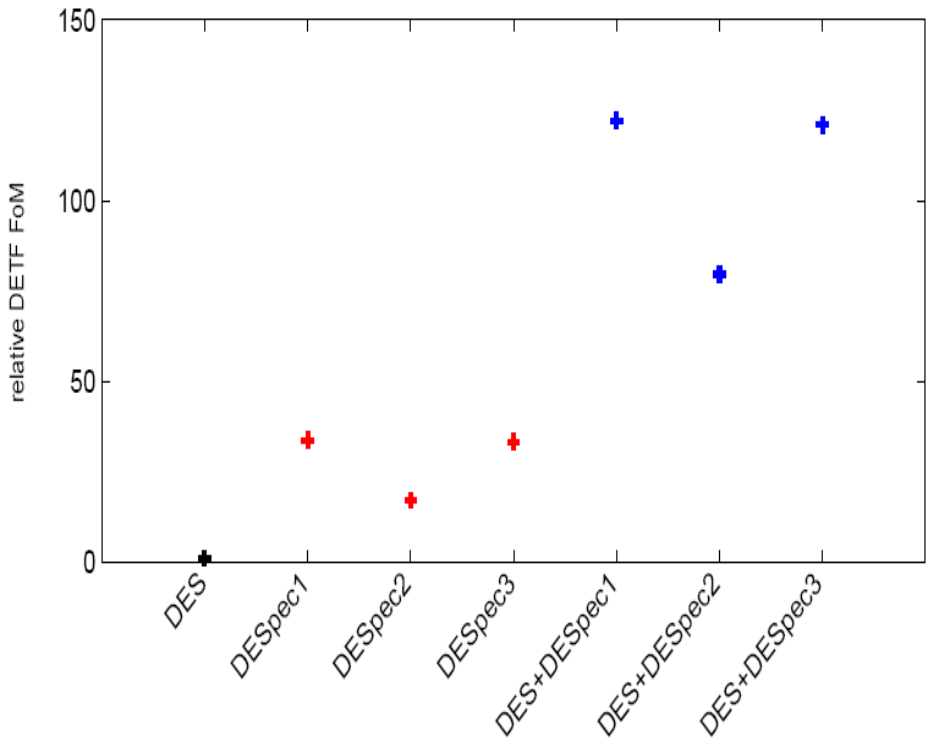

Figure 5.5: DETF Figure of Merit normalised to DES weak lensing alone (black cross). Red crosses show the FoMs achieved by the DESpec RSD measurement alone for each of the three proposed DESpec survey strategies. Blue crosses show joint constraints from DES WGL only and the DESpec RSD information. Joint constraints assume the two datasets are completely independent. All results have been marginalised over the remaining cosmological parameters and galaxy bias.

The WGL, RSD, WGL+RSD analysis was carried out for each of the DESpec survey strategies mentioned above. The results are shown in Fig. 5.5 giving the DETF FoM for each survey and combination. We can see that the general trend is repeated for all surveys- DESpec RSDs are significantly more constraining than DES WGL alone and the joint constraints are several times stronger again. Survey strategy B, the magnitude limited sample, is the weakest of the three strategies. This may be connected to the low fraction of high-z galaxies, limiting the constraints it can place on the time-evolution of w.

5.4 Modified Gravity and spec-z surveys

The results we have presented are preliminary but striking and provide a strong indication of the usefulness of combining a DES-type photo-z survey with a follow-up spectroscopic survey. We

must, however, be wary not to base our survey priorities on the pursuit of a single number. This is a common critique of the DETF FoM, useful though it is. Optimising a survey or multi-survey programme to achieve the best value of a metric for a single parameterisation of dark energy may miss some elements of the bigger picture.

We have already discussed the importance of continuing to test General Relativity, our standard model of gravity by attempting to constrain deviations from GR. We suggest that this may be a useful complement to the DETF approach when considering the interplay between DES WGL and a spec-z follow-up.

We measure deviations from GR using the same parameterisation as chapters 3 & 4. We enlarge our parameter space, including two MG parameters, Q_0 and R_0. These modify the Poisson equation and the ratio of metric potentials respectively as:

$$k^2\Phi = -4\pi G Q(z) a^2 \rho \tag{5.7}$$

$$\Psi = R(z)\Phi \tag{5.8}$$

where for practicality we set $Q(z) = (Q_0 - 1)a^s$, $R(z) = (R_0 - 1)a^s$ and $s = 3$. As discussed in the previous chapters, $s = 3$ is a sensible fiducial model because the laws of gravity are expected to approach General Relativity at high redshift because the density of the universe is greater. As before, we assume that our MG model mimics the fiducial expansion history of a dark energy component with $w_0 = -1$, $w_a = 0$. We assume broad, flat priors on Q_0 and R_0.

First we replicate the constraint on MG using DES WGL alone, and DES WGL assuming all redshifts are perfectly known. The results are shown in Fig. 5.6 and produce similar results to the DE constraints. Spectroscopic redshift knowledge improves the constraining power of cosmic shear but not by a significant amount.

We begin to see a very different picture when we contrast the DES WGL and DESpec RSDs in the MG domain. From Fig. 5.7 we see that, in strong contrast to the dark energy results, the DES WGL constraints (blue contours) are significantly stronger than those from DESpec RSDs (green contours). Very obvious from the plot is the fact that the degeneracy direction in the MG plane is very different, if not quite orthogonal, for the two probes. We may have expected a result qualitatively similar to this from the fact that WGL and RSDs respond differently to the metric potentials. Cosmic shear, being an effect of gravitational distortion on light, repsonds to the sum of the metric potentials, $\Psi + \Phi$. In contrast, RSDs depend upon galaxy position and peculiar velocity which are sourced by the Newtonian potential alone, Ψ. The different dependencies produce different degeneracy directions.

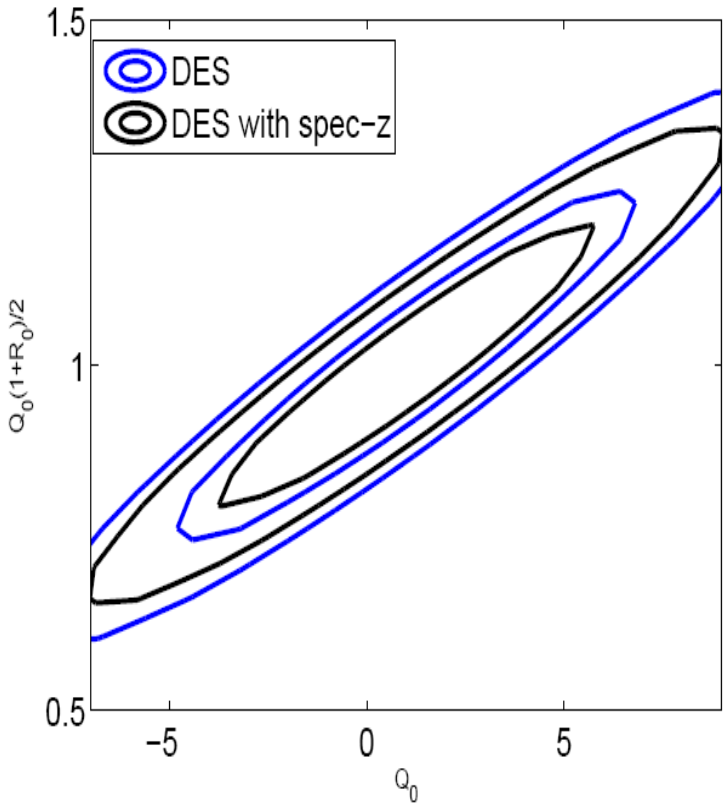

Figure 5.6: Constraints on the modified gravity parameters, Q_0 and R_0, marginalised over cosmological parameters and nuisance parameters. The contours show 68% and 95% confidence limits respectively. Blue contours show the constraints from the standard DES survey set-up with photometrix redshift error $\delta_z = 0.07$. Black contours show constraints for the standard DES survey as if all redshifts were known precisely, $\delta_z = 0.0001$.

Even more encouraging is the fact that the joint constraints (red contours) are significantly more constraining than a simple sum of the two sets of contours would produce. The joint constraint is $\sim$2 orders of magnitude tighter than the DES WGL constraint alone using our MG FoM defined in chapter 3. As in the dark energy case, the combination of probes is breaking parameter degeneracies in the parameter space which we are marginalising over, producing vastly superior constraints to those from either probe alone.

Fig. 5.8 summarises our MG results, showing MG FoM for the different probe and DESpec survey strategies. The results of the constant number density survey are broadly similar to those using the more complex survey strategies. In each case RSD distortions are significantly less constraining than WGL, while the joint constraints are $\sim$ two orders of magnitude tighter than DES WGL alone. It is clear that the synergy between cosmic shear and RSD provides a very

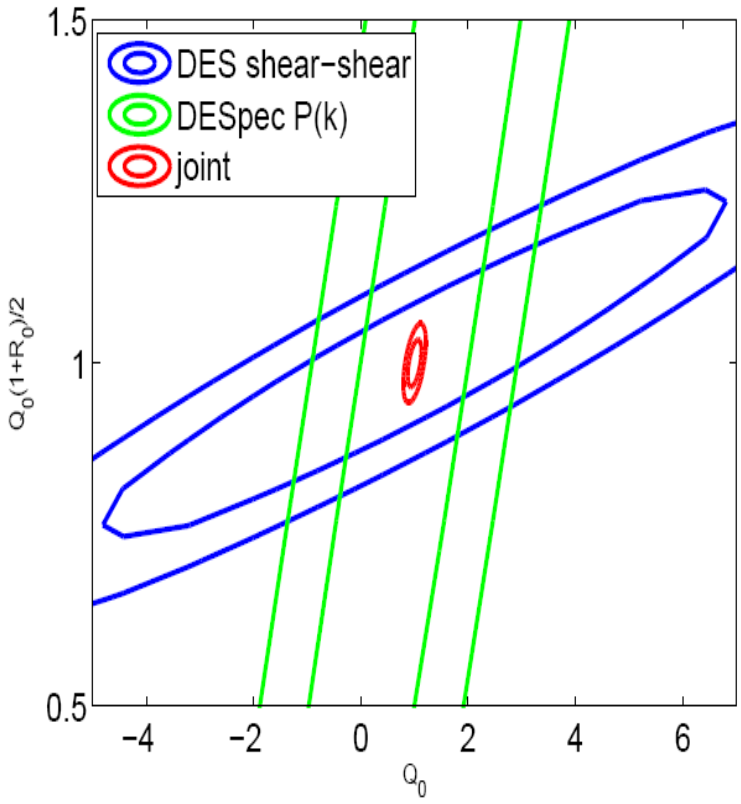

Figure 5.7: Constraints on MG parameters, marginalised over cosmology and galaxy bias nuisance parameters. Inner and outer contours are 68% and 95% confidence limits respectively. Blue contours show constraints for DES WGL alone. Green contours are DESpec RSDs and red contours are the joint constraints, assuming WGL+RSD can be added independently.

powerful approach to constraining deviations from GR.

5.5 Cross-correlations

In the preceding analysis we combined the DES WGL and DESpec RSDs results as if they were independent, simply summing their respective Fisher Matrices. This would be an acceptable approach if we were to combine DES with a RSD survey from a completely separate path of the sky, for example the northern hemisphere BOSS survey. The results are clearly impressive and show that a DESpec-type follow up has some real science potential but the results presented are generic for any spectroscopic survey with similar survey strategy.

One of the key features of the DESpec proposal is that it would be a follow-up survey on the same patch of the sky as DES. Before we can fully understand the impact of putative DESpec data

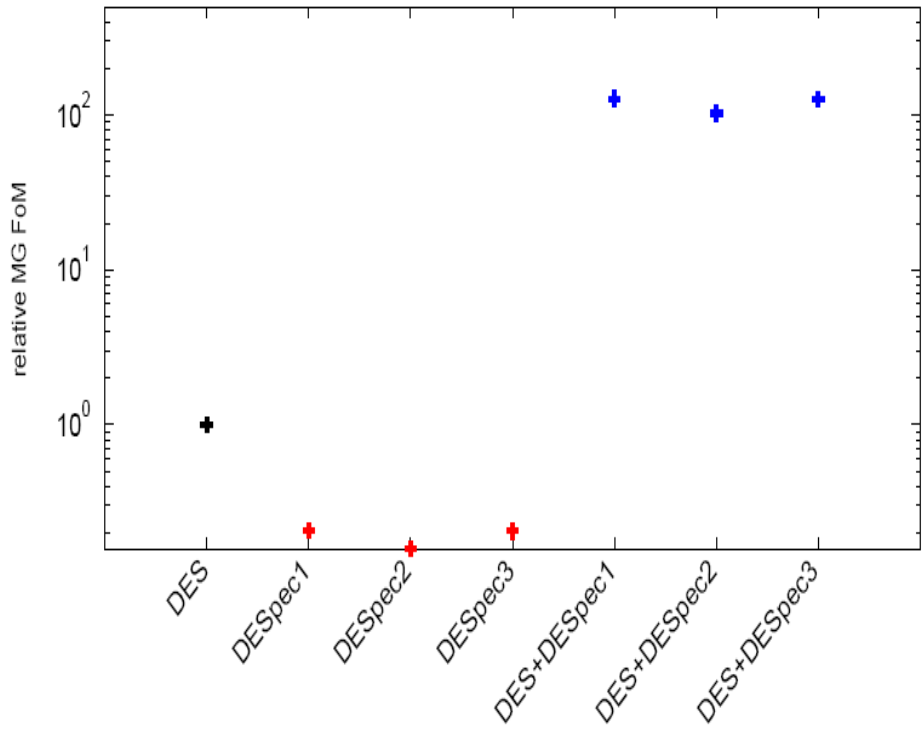

Figure 5.8: MG Figure of Merit normalised to DES weak lensing alone (black cross). Red crosses show the FoMs achieved by the DESpec RSD measurement alone for each of the three proposed DESpec survey strategies. Blue crosses show joint constraints from DES WGL only and the DESpec RSD information. Joint constraints assume the two datasets are completely independent. All results have been marginalised over the cosmological parameters and galaxy bias. Note the log-axis.

on constraints from DES we will have to take the overlapping survey areas and source catalogues into account. In practical terms this involves populating the off-diagonal covariance matrix for a joint data vector of cosmic shear and galaxy/velocity power spectra. To enable this, it may be useful to re-express our RSD power spectra as $C_l s$, including the Kaiser effect as a correction term.

That there will be some significant "cross-talk" between cosmic shear and RSDs is clear on the level of basic physics. A spec-z survey will be tracing the galaxy positions and peculiar velocities (through the RSDs) of a population of foreground galaxies. These observables respond to the background matter density field at foreground redshifts. This density field is, in turn, responsible for gravitationally lensing the light from more distant background sources in our photo-z survey. This correlation is an important statistical effect that must be accounted for. The proper treatment

of correlations across redshifts has, in the past, lead to the discovery of new physical processes such as the Integrated Sachs-Wolfe (ISW) effect which may act as new probes of cosmology.

A complicating factor which we have not considered is the possible double-counting of spectroscopic redshift information. Most likely a non-DESpec source will be used to calibrate the DES photometric redshifts. However if we were to consider using DESpec for this and then sought to combine these constraints with RSDs measured directly from DESpec, it would constitute an unacceptable double use of the spec-z data. A full study of multiple probes and self-calibration techniques is required to optimise the joint analysis strategy of photo-z and spec-z information.

CHAPTER 6

IMAGE SIMULATIONS FOR GREAT08 AND DES

I have worked on practicalities towards getting the shear power spectra from data, as used in the previous chapters. I produced all the galaxy catalogues for the GREAT08 shear measurement challenge and have tested the Dark Energy Survey data portal and shear catalogues.

6.1 GREAT08

6.1.1 Introduction

The use of weak gravitational lensing (WGL) as a cosmological tool relies on our ability to accurately measure the shape of galaxies and convert shape measurements into measurements of the induced cosmic shear from distortions due to the mass distribution in the Universe. The problem of shear measurement is a difficult one and many groups have proposed competing methods since the emergence of cosmic shear as a viable probe just over a decade ago.

One approach that the weak gravitational lensing community has adopted to test shear estimation methods and benchmark them against one another, is to conduct "challenges" where groups compete to analyse simulated sheared galaxy images and extract the most accurate estimate of the shear. The first comprehensive shear estimation challenges went under the name Shear TEsting Programme (STEP) (Heymans *et al.* (2006), Massey *et al.* (2007)). STEP went through four main iterations between 2005 and 2007, each composed of increasingly realistic galaxy simulations. The programme was aimed at existing members of the cosmic shear community.

GREAT08 was conceived as a relaunch of the shear measurement challenge idea. For the first time the challenge was aimed at computer scientists, as well as the weak lensing community. Dr. Sarah Bridle at UCL lead the collaboration responsible for designing, making, distributing and advertising the challenge simulations. My involvement included input into the challenge design, production of all the simulation catalogues, assistance and quality control for the creation of the simulated images themselves, as well as promotion of the challenge at international conferences and the collation and analysis of the challenge results.

In this section I will review the GREAT08 concept and the challenge of shear measurement; discuss the design and creation of the challenge images; summarise the competing methods and challenge results and discuss future plans including the GREAT10 challenge. This work draws on the GREAT08 Challenge Handbook (Bridle *et al.* 2009), the GREAT08 results paper (Bridle *et al.* 2010*b*) and the GREAT10 Challenge Handbook (Kitching *et al.* 2010), all of which I helped to write.

6.1.2 The Backward Process

Measuring cosmic shear from galaxy images is far from trivial. The induced ellipticity from cosmic shear is very small, $\gamma \sim 0.01$. On top of this, galaxies themselves are not perfectly circular, they have their own intrinsic ellipticity, typically an order of magnitude or more higher than the shear signal. Of course we have no access to the unlensed galaxy image so we must find statistical ways to disentangle the induced shear, which can tell us about cosmology, from the measured shear of a particular galaxy which combines the shear part with a much larger intrinsic ellipticity.

Shear measurement is an example of an *inverse problem* as illustrated in Figs 6.1 and 6.2. The forward process describes what happens to the light which forms the galaxy image between emission and the observer. The initially unlensed galaxy image is distorted by the gravitational potential of large scale structure as it propagates through the Universe. It acquires an additional cosmic shear and a (de)magnification (which we ignore for the rest of this chapter). Upon reaching the observer the light passes through the optical system of the telescope (as well as the turbulent atmosphere if the telescope is ground-based). This will inevitably induce distortions in the image which we commonly describe by convolution with a kernel called the Point Spread Function (PSF).

There are a number of additional contributions which affect the image including a roughly constant background source level due to sky emission and detector noise. Light is collected on CCD chips similar to, if larger and more sensitive, than those in a digital camera. The chips contain

detector elements, or light buckets, called pixels so the resulting image is *pixelised*, limiting the possible resolution. There is an inevitable amount of noise in the image, Gaussian noise comes from detector effects but there is also Poisson, or shot, noise from the finite number of photons arriving in each pixel.

In practice any attempt to measure cosmic shear from galaxy images will record the type of images shown on the right hand of Fig. 6.1- noisy, pixellated and distorted. The extraction of the shear information is therefore a type of inverse process, illustrated in Fig. 6.2, where we try to reconstruct the shear field by reversing all the distorting processes that have occured during measurement.

In addition we have noted that, as observers, we have no access to the intrinsic, unlensed galaxy shape. In fact the situation is even worse as we do not even have a particularly good understanding of the general distribution of galaxy shapes. Traditionally galaxies were described, according to the Hubble galaxy morphology scale, as being spiral, elliptical or irregular. In reality many galaxies are somewhere in between these classifications. Further work in the study of galaxy formation and increased numbers of low redshift, low noise imaging surveys will help our understanding of galaxy intrinsic shape (Voigt & Bridle 2010*a*; Bernstein 2010).

In the next sections we briefly summarise some of the important steps in any general reconstruction from noisy images.

Star-Galaxy Separation

An optical measurement of fields on the sky will detect not only the small faint galaxies that we are interested in but also small faint stars from within our own galaxy. The first step in any pipeline to measure cosmic shear is to identify stars and galaxies and "cut-out" a postage stamp around them so that we can work with each object individually. This last task can be complicated when different galaxies have overlapping isophotes.

Incorrectly identifying stars from a sample can bias measured cosmic shear results. Stars in our galaxy are distant enough to appear as point sources. Even so, they will be convolved with the PSF kernel due to seeing and telescope optics. We can therefore use the shapes of stars to reconstruct this PSF. However if we wrongly include some galaxies in our star sample then we will incorrectly reconstruct the PSF, obtaining a PSF which is too big due to the extra convolution by the rogue galaxies. The PSF is roughly equal to the size of our target galaxies (2-3 pixels across), so getting the PSF wrong has a big impact on the measured shear. If we use a PSF which is too big then the reconstructed galaxies will be too small and have to be more elliptical to match the

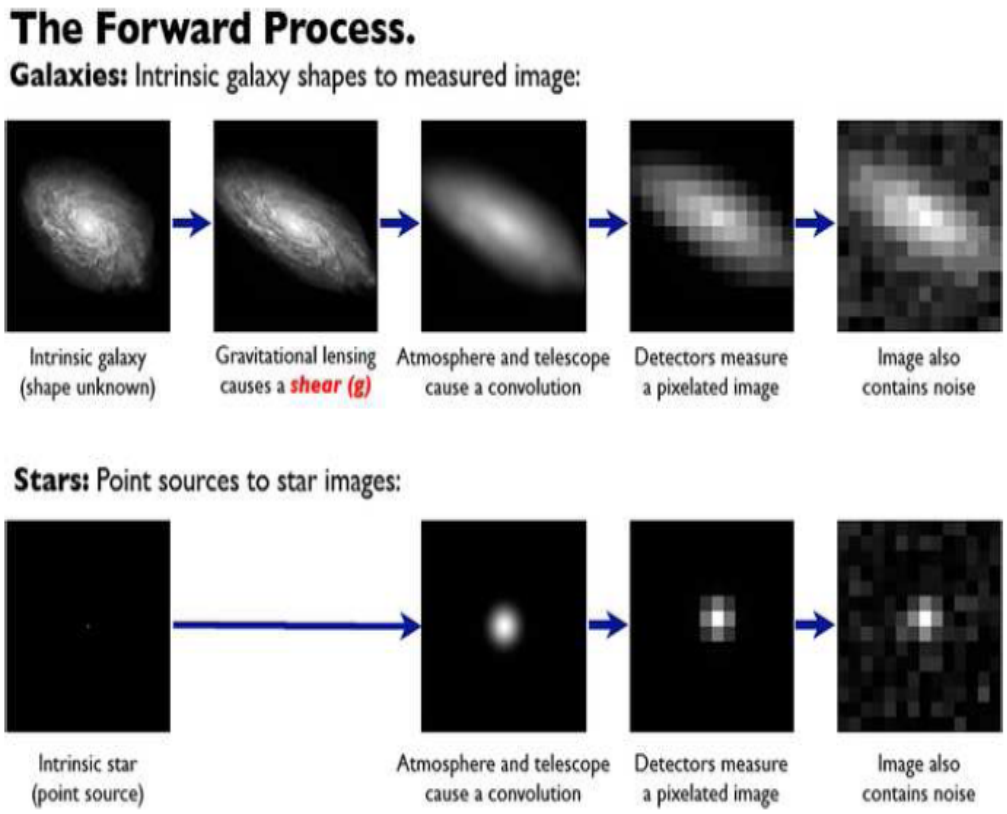

Figure 6.1: Illustration of the forward problem. Upper panels show how the original galaxy image is sheared, blurred, pixelised and made noisy. Lower panels show the equivalent process for (point-like) stars. We only have access to the most right-hand images. Bridle *et al.* (2009)

observed ellipticity of the convolved image, after convolution by our wrong enlarged PSF. This will overestimate the shear.

PSF Measurement & Interpolation

Point-source stars give us the opportunity to measure the PSF. Generally the PSF will vary across the telescope's field of view and hence across the image. A sufficient sampling of bright stars across the image allows us to measure the PSF at different points then interpolate between them to model the PSF across the whole image. Various methods have been considered to characterise the PSF, from fitting a simple polynomial to decomposing it into principal components (Jarvis & Jain 2006).

The point of measuring the PSF is so that we can correct for its effects in the images of galaxies we are interested in. This is a difficult but extremely important step in the analysis and has been

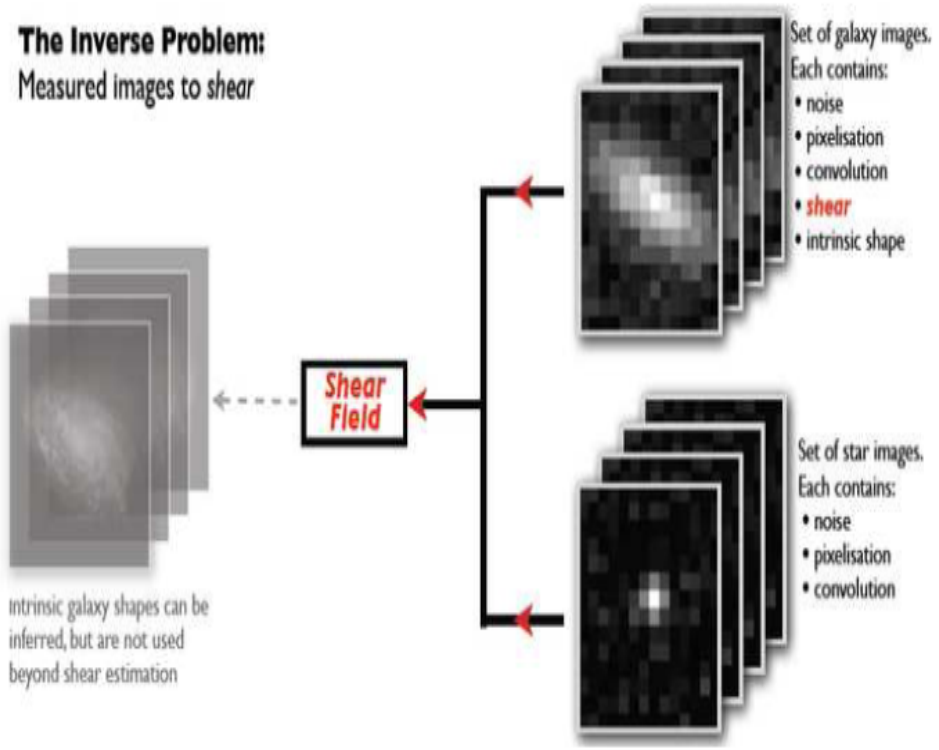

Figure 6.2: Illustration of the inverse problem. We begin on the right with a set of galaxy and star images. The full inverse problem would be to derive both the shears and the intrinsic galaxy shapes. However shear is the quantity of interest for cosmologists. Bridle *et al.* (2009)

tackled in different ways by different groups. A simple approach is to subtract the effects of the measured kernel from our measured quantites such as shear (Kaiser, Squires & Broadhurst 1995, hereafter KSB). A more advanced technique is to characterise the PSF in terms of basis functions then forward fit to produce a deconvolved model of each measured galaxy (e.g. im2shape, see Bridle et al. 1999, LensFit see Miller *et al.* (2007)). The deconvolved images are then used to measure quantities of interest, free from PSF effects.

Correcting for Pixelisation

All modern astronomical telescopes collect light on Charged-Coupled Devices (CCDs) which use wafers of silicon arranged in pixels to collect photons and convert them into electrons which can be read out electronically. The effects of pixelisation must be accounted for in any realistic shear measurement method. Quadrupole moments methods such as KSB do not usually have any

mechanism for accounting for pixelisation. In principle methods which fit a model to a galaxy shape can simply deconvolve the measured image with square top-hat functions equivalent to the square pixel scale.

An additional effect has been noted due to the physics of reading out electrons from each CCD pixel (Massey *et al.* 2010). Particularly acute in space, where telescope optics are more likely to be damaged by cosmic rays etc., electron readout from each pixel may not be perfect and can create smears across pixels. The fact that we are trying to measure galaxy ellipticity makes this of high concern.

Countering Noise

Our galaxy images will be subject to noise from our imperfect telescope optics and CCD detectors (Gaussian) and shot-noise on the image from the finite number of photons collected (Poisson). In addition, our ignorance of the intrinsic ellipticity of a galaxy contributes "shape noise" to our measurements which can be beaten down by averaging the ellipticites of many galaxies.

6.2 GREAT08 Concept

We have mentioned that the GRavitational lEnsing Accuracy Testing 2008 (GREAT08) challenge was envisaged as a reboot of the shear estimation challenge programme after the end of the STEP 4 challenge. By the fourth iteration the STEP simulations had become relatively complicated and it was becoming difficult to distinguish which features of the simulated galaxies were proving limiting factors to the accuracy of shear estimation methods (Massey *et al.* 2007*b*).

GREAT08 adopted a back-to-basics approach, producing simpler simulated images than the final STEP image suites in an effort to better distinguish differences in method performance. The orientation of the GREAT08 challenge was also very different to previous efforts. As well as being advertised within the weak lensing community and in cosmology generally, GREAT08 succesfully applied for status as a PASCAL network challenge. This is a European based network of computer scientists who regularly run challenges to stimulate activity in the fields of data mining, pattern recognition and image analysis. By opening the problem of shear measurement outwards and posing it as a statistical challenge, it was hoped that the cosmic shear community could capitalise on useful innovations from widely different fields of study as well as forge links with members of other academic disciplines whose work was complementary to some of the challenges facing today's cosmologists.

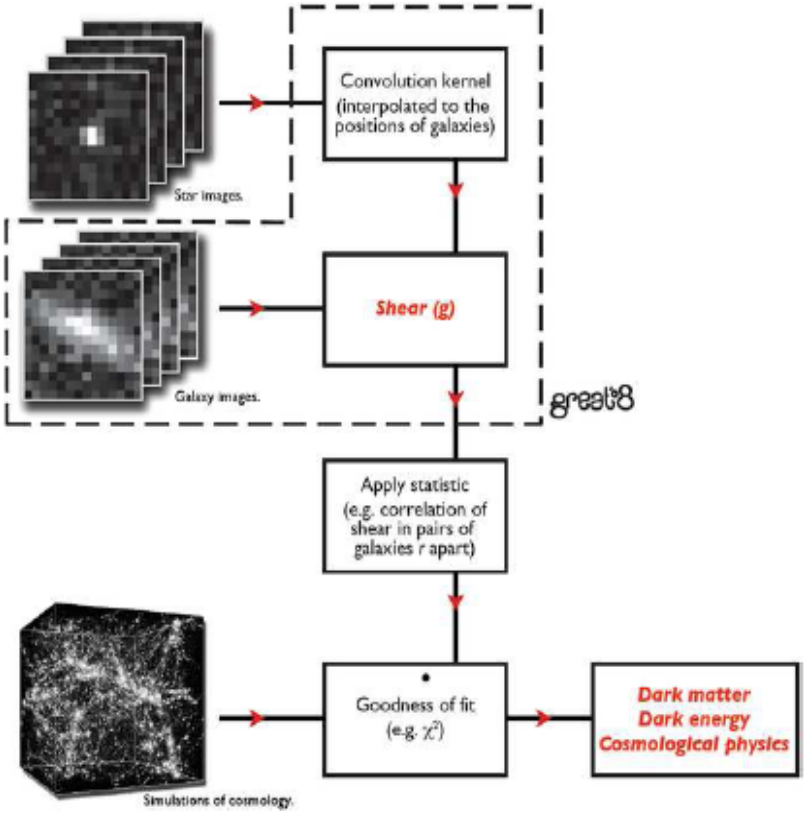

Figure 6.3: Flowchart showing the extent of a full conventional cosmic gravitational lensing data analysis pipeline, from measurement of the convolution kernel using the shapes of stars to measurement of cosmological parameters. The GREAT08 Challenge focuses on the steps enclosed in the dashed box. (Bridle *et al.* 2009) and Andrey Kravstov (Uni. of Chicago)

The GREAT08 challenge comprised a set of simulations and a handbook explaining the problem of shear estimation phrased as a statistical problem and noting some of the more successful attempts that had been made to tackle it in the past. A website [1] provided more support, documentation and a forum to discuss shear measurement issues as well as a portal through which competing groups and individuals could submit their results- the shear values they measure from the simulated galaxy images. Fig. 6.3 shows the shear measurement pipeline and highlights the area which the GREAT08 challenge was designed to test. The next section discusses the design and creation of the simulations in detail.

[1] `http://www.great08challenge.info/`

Table 6.1: Parameters for the LowNoise_Known simulations. R_{gp}/R_p is the ratio of PSF convolved galaxy Full Width at Half Maximum (FWHM) to the PSF FWHM. 'b or d' describes the fact that 50% of the galaxies in each set have de Vaucouleurs profiles (bulge only) and 50% have exponential profiles (disk only). The parameters for LowNoise_Blind are the same except the galaxies are a mix of the two components as described in the text. The parameters for RealNoise_Known are the same as for LowNoise_Known except the SNR is 20.

	Fiducial	Lower value	Upper value
SNR	200	N/A	N/A
R_{gp}/R_p	1.4	1.22	1.6
PSF type	Fid	N/A	N/A
Galaxy type	b or d	N/A	N/A

6.2.1 GREAT08 Simulations

The GREAT08 images are provided in sets of 10,000 objects in a single `FITS` file. Each object is generated on its own grid of 39×39 pixels and these postage stamps are patched together for convenience in a 100×100 layout, with a 1 pixel border, thus each set is a patchwork image of 4000x4000 pixels. Each galaxy postage stamp is generated using the following sequence: (i) simulate a galaxy model; (ii) convolve it with a kernel, referred to as the point-spread function (PSF); (iii) bin up the light in pixels; and (iv) apply the noise model. The PSFs used are given in table 6.2.2. Each postage stamp is produced using a list of parameters specifying the individual object and simulation properties. We describe the catalogues of these properties in 6.2.3. The method used to produce images from the catalogues is overviewed below. Example images are shown in Fig. 6.4.

Four different groups of galaxy images were provided in GREAT08: (i) low noise galaxy images for which the true shears were provided during the Challenge, labelled LowNoise_Known; (ii) low noise galaxy images for which there was a blind challenge to extract the true shears, labelled LowNoise_Blind; (iii) realistic noise galaxy images for which the true shears were provided, labelled RealNoise_Known; and (iv) realistic noise galaxy images with blind shear values, RealNoise_Blind. This RealNoise_Blind group formed the main GREAT08 Challenge. The Low Noise sets were provided to determine the effect of realistic image noise on shear estimation methods while the Known sets were provided to act as training sets, particularly useful for neural network methods.

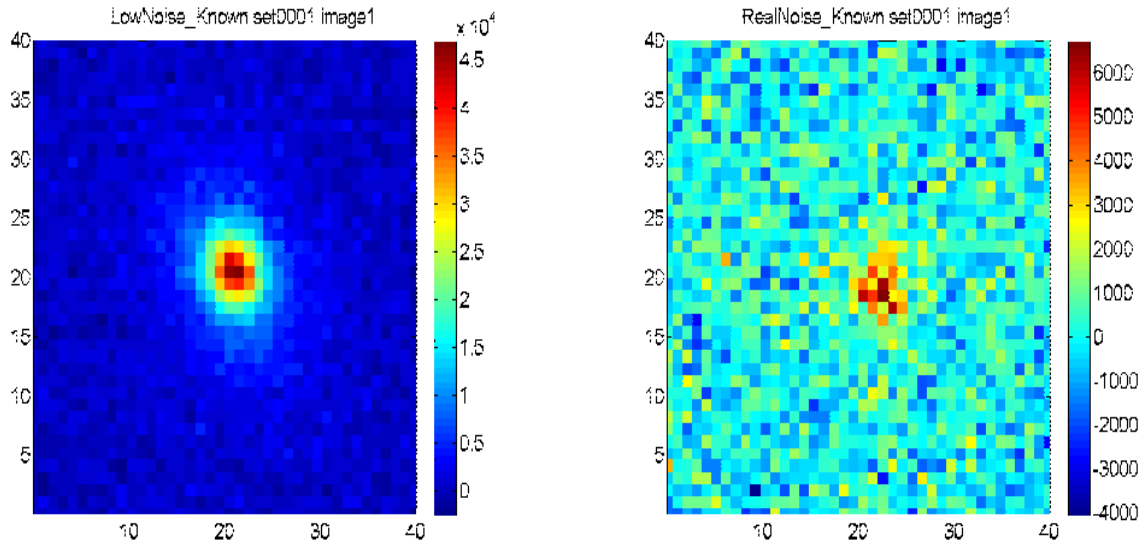

Figure 6.4: Left: The first galaxy of the first LowNoise_Known FITS image. Right: The first galaxy of the first RealNoise_Known FITS image. The signal is a factor of ten smaller for the RealNoise images than the LowNoise images, making the problem much more challenging.

The parameters for each set in LowNoise_Known were determined using Table 6.1 and the parameters for each set in RealNoise_Blind were determined using the lower panel of Table 6.2. There is a range in SNR, galaxy size, PSF ellipticity and galaxy type. One branch of the RealNoise_Blind holds all parameters at their fiducial values. Each of the 4 variable parameters has a 'lower' and an 'upper' value relative to the fiducial. When each of these values is used all other parameters are fixed at the fiducial values. This makes 9 different branches in total. In each branch there are 6 realisations of each of 50 different shear values, making 2700 sets with 10,000 galaxies in each.

Images are generated by sampling from the galaxy light distribution, sampling from the PSF,

Table 6.2: Parameters for the RealNoise_Blind simulations. The PSF models and other parameters are defined in detail in 6.2.2 and 6.2.3.

	Fiducial	Lower value	Upper value
SNR	20	10	40
R_{gp}/R_p	1.4	1.22	1.6
PSF type	Fid	Fid rotated	Fid $e\times$ 2
Galaxy type	b+d	b or d	b+d offcenter

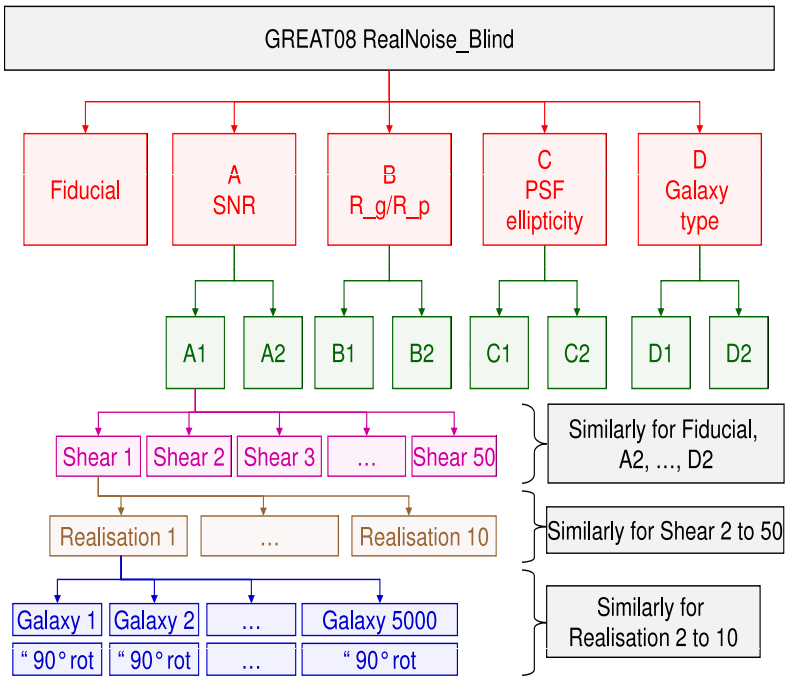

Figure 6.5: Schematic of the galaxy parameters used in RealNoise_Blind. Each realisation corresponds to a different set or `FITS` image file containing 10,000 galaxies. The bottom row of boxes represents galaxies with the same properties as the penultimate row of boxes, but rotated by 90 degrees. (Bridle *et al.* 2010*b*)

adding the sample positions to simulate convolution, binning the samples onto a pixel grid, and then applying the noise model. In brief, samples are first generated from the circular galaxy profile. Next, they are stretched to have the required ellipticity and then sheared. Samples are then drawn from the circular PSF distribution and made elliptical using the shear distortion equations. The x and y coordinate values of each galaxy sample are added to those of the PSF sample, to simulate convolution, and finally the samples are binned into pixels. More details of this process are given below.

6.2.2 PSF models

In an attempt to isolate problems in the shear estimation pipelines and make the challenge more accessible we provided maximal information about the PSFs used during the competition.

Table 6.3: PSF ellipticities.

PSF type	Filename	e_1	e_2
Fid	set0001	-0.019	-0.007
Fid rotated	set0002	0.007	-0.019
Fid $e \times 2$	set0003	-0.038	-0.014

The PSFs had a truncated Moffat profile

$$I_p(r) = \begin{cases} \left(1 + \left(\frac{r}{r_d}\right)^2\right)^{-\beta} & r < r_c \\ 0 & r >= r_c \end{cases} \tag{6.1}$$

where we set $\beta = 3$. This profile is motivated by the combination of diffraction limited optics with random Gaussian blurring by the atmosphere and is therefore reasonably representative of PSFs for ground-based telescopes. The scale radius r_d was determined by setting the Full Width at Half Maximum (FWHM) to 2.85 pixels. r_c was set to twice the FWHM. Three different PSFs were used in the GREAT08 Challenge, each with a different ellipticity, as shown in Table 6.3.

Star catalogues consisted simply of the position of the point source. The convolution kernel and image generation method are described below.

6.2.3 Galaxy Catalogue Generation

The information provided in this subsection was not available during the Challenge.

In general, the galaxies in GREAT08 are the sum of two components, each with a Sersic (Sersic 1968) intensity profile

$$I(r) = \begin{cases} I_o \exp\left(-\kappa (r/r_e)^{1/n}\right) & r < 4r_e \\ 0 & r >= 4r_e \end{cases} \tag{6.2}$$

where $I(r)$ is the amount of light per unit area at a radius r, and $\kappa \simeq 2n - 0.331$ (Peng *et al.* 2002). The scale radius r_e and the total intensity (which determines I_o) are free parameters specified in the catalogues. The first component, with $n = 4$, is an approximation to the central bulge component of galaxies, corresponding to a de Vaucouleurs profile. The second component, with $n = 1$, is an approximation to the exponential disk component of galaxies. Circular galaxy images are made according to the profile $I(r)$ described above and then distorted according to the galaxy ellipticity and shear as described below.

Table 6.4: Galaxy scale radius values for single-component galaxy models. The left hand column gives the ratio of PSF convolved galaxy FWHM to the PSF FWHM. The middle column gives the scale radius for a single component disk model. The right hand column gives the scale radius for a single component bulge model. These values are interpolated to produce scale radius values for two-component models, as described in the text.

R_{gp}/R_p	Disk $r_{e,d0}$	Bulge $r_{e,b0}$
1.22	0.82	1.59
1.4	1.3	3.8
1.6	2.4	18.0

Therefore the objects are circularly symmetric before the intrinsic ellipticity and gravitational shear is applied. More realistic galaxies will be used in future GREAT Challenges. This was investigated in the precursor Shear TEsting Programme (STEP), for which simple galaxies like those in GREAT08 were used for STEP1 (Heymans *et al.* 2006) and sophisticated shapelet galaxies based on the Hubble Deep Field were used in most of the STEP2 simulations (Massey *et al.* 2007*b*). However, note that after the intrinsic ellipticity is applied the galaxies no longer have elliptical isophotes. It was shown that this makes the shear measurement problem non-trivial (Lewis 2009) at the kind of accuracy level required in GREAT08 (Voigt & Bridle 2010*b*).

The x and y positions of the bulge component were each drawn from a Gaussian of standard deviation 1.2 pixels centred on the middle of the postage stamp.

For each object, the total flux (integral of $I(r)$ over the postage stamp) in the disk component, as a fraction of the total flux in both components, is in general a random number drawn from a uniform distribution between 0 and 1.

The scale radii r_e of each component were set by considering high resolution circular galaxy images after convolution with the appropriate PSF. For single-component models (i.e. when the bulge to total flux is zero or unity), r_e is set such that the convolved image has a FWHM of 1.4 times that of the PSF, $F_{gp} = 1.4F_p$, in the fiducial branch. Values 1.22 or 1.6 were used for some other branches to explore the effect of galaxy size, as detailed below (see Tables 6.1 and 6.2).

The ellipticities of the bulge and disk were drawn from

$$P(\epsilon) \quad = \quad \epsilon \left(\cos\left(\frac{\pi\epsilon}{2}\right)\right)^2 \exp\left(-\left(\frac{2\epsilon}{B}\right)^C\right) \tag{6.3}$$

with $B = 0.05$, $C = 0.58$ for the bulge and $B = 0.19$, $C = 0.58$ for the disk; $\epsilon \equiv (a^2 - b^2)/(a^2 + b^2)$ where a and b are the major and minor axes respectively. See Fig. 6.6 for a plot of

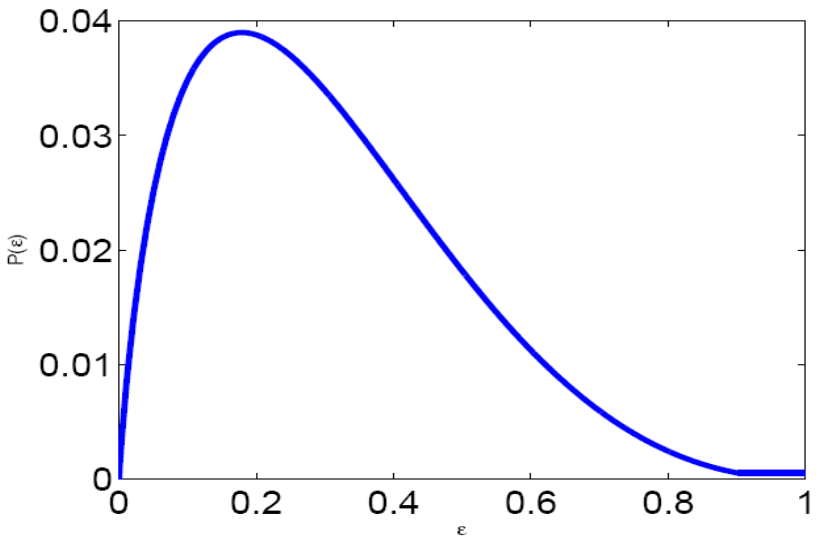

Figure 6.6: Distribution of galaxy ellipticities as used in the GREAT08 competition and defined by eqn. 6.3.

this distribution. Since ellipticities close to unity become unphysical, we truncate the distribution at $\epsilon = 0.9$ and set all objects with $\epsilon > 0.9$ to have $\epsilon = 0.9$. This distribution was loosely motivated by results from the APM survey (Crittenden *et al.* 2001); The bulge and disk ellipticities are drawn independently from the above distributions and are thus uncorrelated. The angle between the bulge major axis and the positive x axis is drawn from a uniform distribution between 0 and 180 degrees. The disk angle is equal to the bulge angle but perturbed by a Gaussian of standard deviation 20 degrees.

Five thousand galaxy parameters were simulated per image set by drawing from the above distributions. To minimise noise the parameters were all rotated by 90 degrees to produce the remaining 5000 galaxy parameters. (i.e. all angles are increased by 90 degrees, x positions become y positions, and y positions become negative x positions.) The list was randomised to hide the pairings. This paired rotation was introduced in STEP2 to reduce shape noise. In the absence of a PSF or shear, the shear estimates from each galaxy in a pair are expected to cancel, thus removing noise arising from the intrinsic ellipticities of galaxies.

Signal-to-Noise Ratios (SNR) are assigned in the catalogues and are used during image simulation to set the flux in the galaxy image. For LowNoise images the value is 200, and for RealNoise images the default value is 20, with variations to 10 and 40 within RealNoise_Blind. The definition of this number in terms of the noise model is described in the following subsection.

For LowNoise_Known and RealNoise_Known the galaxies all have just a single component and within each set, each galaxy is assigned a de Vaucouleurs or an exponential profile at random. The galaxies in LowNoise_Blind all have a bulge plus disk two-component model as described in the text above. The majority of the galaxies in RealNoise_Blind have the same two-component model as in LowNoise_Blind. One of the nine RealNoise_Blind branches has single-component galaxies as in the Known simulations. The two-component models all share the same centroid for the bulge and disk, except for one of the nine RealNoise_Blind branches, in which the bulge is off-centred from the disk by a Gaussian of standard deviation 0.3 pixels.

The true shears for LowNoise_Known and RealNoise_Known were provided throughout the challenge. They are Gaussian distributed with a standard deviation of 0.03 in each of g_1 and g_2, and zero mean. These shears are perturbations around the root values $g_1 = (-1, 0, 1, 0, -1/\sqrt{2}) \times 0.037$ and $g_2 = (0, 0, 0, 1, -1/\sqrt{2}) \times 0.037$ and thus do not have zero mean. This distribution is chosen instead of a Gaussian to improve the uncertainties on linear fits to the output versus true shear. For LowNoise_Blind, one position in shear space is drawn from around each root and there is one set with this shear. For RealNoise_Blind, 50 positions in shear space are drawn from around each root and there are 6 sets with each shear, as illustrated in Fig. 6.5.

6.2.4 Image simulations

The galaxy images are created according to the *forward process* using a Monte Carlo simulation technique. The general idea is that the intensity of a pixel in the image of a galaxy is directly proportional to the number of photons falling into that pixel. The photon count at each point depends on the intensity distribution (the light profile) of the galaxy. Therefore, if we draw random samples (photons) from the theoretical light profile function and then count the number of photons falling in each pixel, we obtain the image of galaxy with the required light profile. The circular light profile thus obtained is then reshaped by applying the necessary transformations to the coordinates of the photons. Since the point-spread function (PSF) can be considered as a probability distribution, a similar method can be used to simulate it. The light profile of the galaxy is convolved with the PSF and finally pixelized into a `FITS` image.

In order to simulate the photons distributed by a Sersic Law, we need to find the cumulative distribution of the density given by Equation 6.2. Taking $r_e = 1$ and substituting $R = kr^{1/n}$, we obtain the cumulative distribution as

$$F(R) = -\frac{\Gamma(2n, R)}{\Gamma(2n)}, \tag{6.4}$$

where n is the Sersic index and $\Gamma(a, x)$ is the incomplete Gamma function. The inverse of the distribution can be approximately calculated by using linear interpolation, given that we have an ordered set of values of $\{R, F(R)\}$ for the range of R (e.g., from 0 to 20).

The circular light profile of the galaxy obtained by the method above is made larger, elliptical and rotated according to the values of scale radius r_e, axis ratio q and angle ϕ respectively. These operations can be represented in the form of matrices as

$$\begin{pmatrix} x_e \\ y_e \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{q}\, r_e} & 0 \\ 0 & \sqrt{q}\, r_e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6.5}$$

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} x_e \\ y_e \end{pmatrix}. \tag{6.6}$$

The shear from the gravitational lensing is applied next. This operation can be written as

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} = \begin{pmatrix} 1+g_1 & g_2 \\ g_2 & 1-g_1 \end{pmatrix} \begin{pmatrix} x_r \\ y_r \end{pmatrix}. \tag{6.7}$$

For computational simplicity, we combine all of the above operations into a single matrix given by

$$\begin{pmatrix} r_e\left((1+g_1)c - g_2 s\right)/\sqrt{q} & r_e\left(g_2 c - (1-g_1)s\right)/\sqrt{q} \\ r_e\sqrt{q}\left((1+g_1)s + g_2 c\right) & r_e\sqrt{q}\left(g_2 s + (1-g_1)c\right) \end{pmatrix} \tag{6.8}$$

where $c \equiv \cos\phi$ and $s \equiv \sin\phi$.

Having obtained the light profile of the galaxy, we move on to create a Moffat PSF and convolve it with the galaxy. Using a similar procedure to that described above for the Sersic profile, we can simulate Moffat PSF given by the Equation 6.1. Each sample from the PSF corresponds to the displacement of the photon when convolved with the galaxy. The circular galaxy can be scaled to the required FWHM and made elliptical by applying the transformation

$$\begin{pmatrix} x_p \\ y_p \end{pmatrix} = \begin{pmatrix} 1-e_1 & -e_2 \\ -e_2 & 1+e_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{6.9}$$

Assuming that the number of samples in the light profile and the PSF are the same, the convolution of the image is accomplished by adding the positions of the galaxy and PSF photons. The image is pixelized by counting the number of photons falling into each pixel of the postage stamp and then it is normalized.

The galaxy images in GREAT08 contain two different light profiles. The final image is created by adding together two images with different light profiles. If I_1 and I_2 represent two galaxy images with different light profiles, the final image $I_{\rm final}$ is created by the equation

$$I_{\rm final} = mI_1 + (1-m)I_2, \tag{6.10}$$

where $0 \leq m \leq 1$ is a multiplication factor. Poisson noise is then added to each pixel according to the SNR.

CCD detectors on ground-based telescopes collect a finite number of photons from both astrophysical objects and atmospheric emission. We mimic this effect by adding a background level $B = 1 \times 10^6$ to each pixel, and drawing a number from a Poisson distribution with a mean equal to the total number (background plus galaxy) in each pixel. For numerical convenience we then subtract B from each pixel. For the RealNoise simulations, this background is much larger than the contribution from the galaxy, so this process is closely approximated by adding a Gaussian random number of standard deviation $\sqrt{B}$ with zero mean.

Before the noise model is applied, the total flux in the galaxy is set using the SNR given in the catalogue, and the background level discussed above. In summary we define SNR as the flux divided by the uncertainty in the flux obtained if the true shape (but not normalisation) of the object is known.

For the purpose of the SNR calculations we approximate the Poisson noise as a Gaussian of standard deviation $\sqrt{B}$ for both LowNoise and RealNoise simulations. We follow the definition

$$\mathrm{SNR} = \frac{F}{\sigma_F} \tag{6.11}$$

where the flux F is the sum of the galaxy counts in each pixel I_i

$$F = \sum_i I_i \tag{6.12}$$

and σ_F is the uncertainty in the flux. In general the uncertainty in the flux depends on the assumptions used to measure it. We make the assumption that the true galaxy shape (profile of counts in all the pixels) is known precisely up to an overall unknown scaling which is proportional to the flux. By considering a χ^2 fit it can then be shown that

$$\sigma_F = \sqrt{B}\frac{F}{\left(\sum_i I_i^2\right)^{0.5}} \tag{6.13}$$

and therefore the flux can be set such that

$$\left(\sum_i I_i^2\right)^{0.5} = \mathrm{SNR}\sqrt{B} \tag{6.14}$$

to produce the desired SNR.

We note that the images produced using the above ellipticities and r_e values give some very elliptical images that extend beyond the 39×39 postage stamp. We do not take into account the fact that there are some objects which don't fit the postage stamp. This was a practical limitation of GREAT08 due to the increased data volume required for larger postage stamps. However real data also suffers from edge effects when galaxies are large, for example due to detector defects and objects which are difficult to model. We therefore felt that it was reasonable to include such limitations in the GREAT08 Challenge. Future GREAT Challenges will not isolate the galaxies onto separate postage stamps and therefore will not have this limitation.

6.2.5 GREAT08 Results

6.2.5.1 GREAT08 Quality Factor

To define the success of each competing method and allow for easy ranking, we devised a single number, Q, which would act as a "Quality Factor". This was computed automatically as groups submitted results via the GREAT08 website. The quality factor was then used to update the leaderboard.

Q takes into account that a successful shear estimation method must both filter the noise effectively from the image and remove the PSF convolution kernel. Q is calculated:

$$Q = \frac{10^{-4}}{\langle(\langle g^m_{ij} - g^t_{ij}\rangle_{j\in k})^2\rangle_{ikl}} \tag{6.15}$$

where g^m_{ij} is the ith component of the measured shear for simulation j, g^t_{ij} is the corresponding true shear component, the inner angle brackets denote an average over sets with similar shear value and observing conditions $j \in k$, and the outer angle brackets denote an average over simulations with different true shears k, observing conditions l and shear components i.

In fact a Q-factor is calculated for each simulation branch and the total value arrived at via

$$\frac{1}{Q} = \frac{1}{\langle Q_l\rangle_l}. \tag{6.16}$$

This definition has the effect of strongly penalising methods that perform poorly in any single simulation branch, which is useful because the simulation branches are all chosen to be realistic scenarios in which we need to be able to measure good shears.

In the past (including STEP) the effectiveness of a method was generally quantified by two numbers, m the multiplicative bias on shear and c, the additive bias, i.e.

$$g^m_i - g^t_i = m_i g^t_i + c_i\,. \tag{6.17}$$

The Q-factor takes into account both sources of error but it can be useful to calculate them separately and learn more about the particular weaknesses/strengths of methods.

Our current shear estimation methods are sufficiently accurate for the quantity and accuracy of current data. Amara & Refregier (2007) show that a deep full-sky (e.g. Euclid-like) survey requires that the additive error $c < 0.0003$ and the multiplicative error $m < 0.001$. For a pure additive error this translates to a requirement that $Q > 1000$ and we set this as our target for GREAT08 because additive errors are much more difficult to self-calibrate using pairs of tomographic redshift bins (Huterer *et al.* 2006) (see also Van Waerbeke *et al.* 2006).

There was an optional procedure for candidates to submit uncertainty estimates alongside their estimated shear values. Methods which have very high error values are flagged as such on the leaderboard even if they achieve a good Q value. In practice a noisy estimator will have significant scatter and find it difficult to achieve a very high Q value.

6.2.5.2 Methods

Here we briefly summarise the types of method which were submitted to the GREAT08 challenge, drawing some general comparisons between them and trying to convey an idea of the landscape of techniques under investigation. Table 6.5 lists the participants. Methods with an asterisk indicate GREAT08 Team entries; these participants had access to the internal details of the GREAT08 Challenge simulations, but they did not consciously use this information in their analyses.

The table schematically describes each method in terms of three action steps. One key dividing line between methods is at what stage the average is performed over galaxy shape. Methods such as HB, AL and USQM average galaxy properties before attempting to fit models, these are known as "stacking" techniques. The advantage of stacking methods is that the individual on-elliptical shapes and intrinsic ellipticity information are averaged out, leaving in some sense a purer shear signal to measure. A criticism of these methods is that they rely on a property of the GREAT08 challenge, the availability of a large number of galaxies all with exactly equal induced shear. This is not the case with realistic galaxy survey images. Intrinsic Alignment effects may hurt stacking techniques particularly hard.

A nuisance of the stacking method is that the galaxies must be lined up on a common centre before stacking to avoid biasing the measured shear. This is known as the centroiding problem. HB's decision to stack Fourier transforms of images avoids this problem entirely by using a version of the galaxy information which is independent of spatial location.

The stacking approach has not been a favoured method for shear estimation within the cosmic

Participant(s)	Key	Action 1	Action 2	Action 3
Hosseini, Bethge	HB	Estimate power spectrum	Average power spectra	Fit sheared circular model ∗ PSF
Lewis	AL	Estimate centroids	Average images	Fit sheared circular model ∗ PSF
Kitching	TK†	Fit sheared circular model ∗ PSF	Combine shear PDFs	Calculate shear
Heymans	CH†	Measure weighted quadrupole moments	Correct for weight and PSF	Average shear estimates
Paulin, Gentile	PG	Fit sheared circular model ∗ PSF		Average shear estimates
Velander	MV	Fit flexed sheared circular model ∗ PSF		Average shear estimates
Kuijken	KK†	Fit sheared circular model ∗ PSF		Average shear estimates
Harmeling, Hirsch, Schölkopf	HHS3	Estimate centroids	Average good images	Fit sheared circular model * PSF
Bridle	SB†	Fit sheared circular model ∗ PSF		Average shear estimates
Harmeling, Hirsch, Schölkopf	HHS2	Estimate centroids	Average images	Fit sheared circula model ∗ PSF
Harmeling, Hirsch, Schölkopf	HHS1	Fit sheared circular Gaussian	Correct for model and PSF	Average shear estimates
Jarvis	MJ†	Fit “elliptical” model ∗ PSF		Average shear estimates
Bridle, Schrabback	USQM†	Measure quadrupole moments - PSF	Average quadrupole moments	Calculate shear

Table 6.5: Table of participants, figure legend identifiers and pseudo-code which attempts to summarise the main actions carried out in each method. “∗ PSF” indicates that a PSF convolved model was fitted. “PDF” stands for probability density function. Daggers after the Key indicate GREAT08 Team entries. More information is provided in the main text.

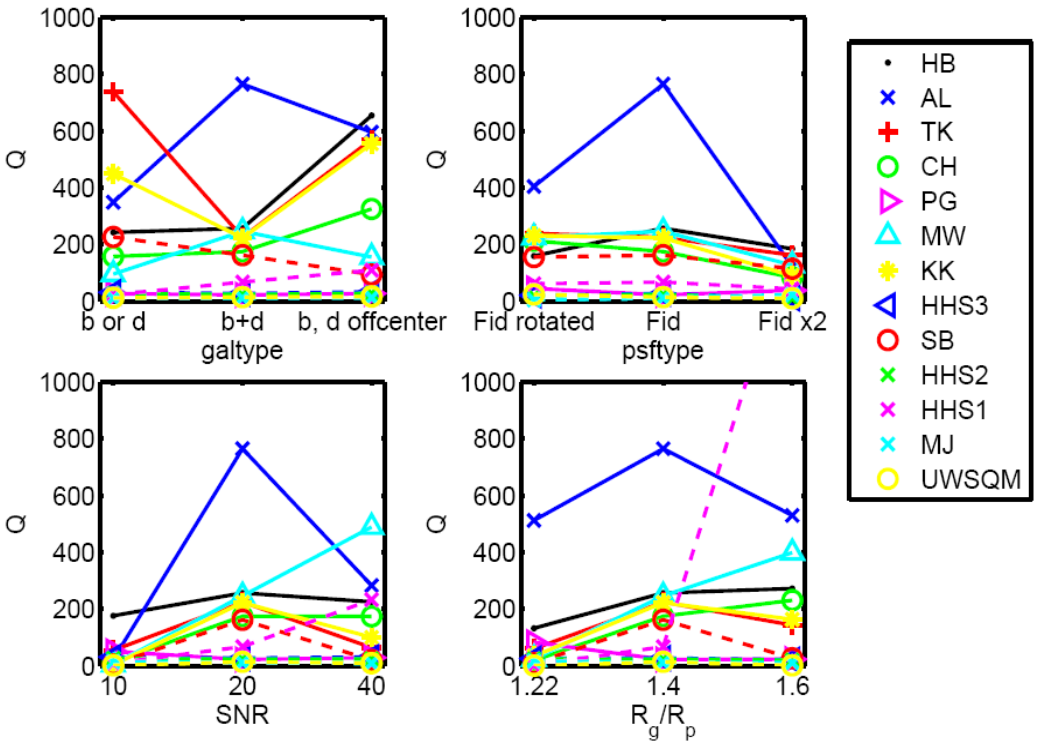

Figure 6.7: Shear measurement figure of merit Q as a function of simulation properties for RealNoise_Blind. (Bridle *et al.* 2010*b*)

shear community. It has been generally assumed that it will prove impossible to stack real galaxy images effectively due to the impact of intrinsic alignments and the variation of shear with angular position. That these methods did so well in the simple GREAT08 challenge has sparked renewed debates about the merits of stacking.

Of the non-stacking techniques some, like CH measure quadrupole moments of each galaxy image. Most other methods involve the fitting of a particular model to each galaxy image, whether the model involves shapelet decomposition, sersic index profiles or simply elliptical Gaussians. The shear estimates for each individual galaxy are then averaged to produce results for each `FITS` file.

6.2.5.3 RealNoise Blind Results

We present a brief summary of the results of the GREAT08 challenge, focusing on the main challenge dataset, RealNoise_Blind.

The winner of the GREAT08 Challenge is clearly the 'CVN Fourier' method by R. Hosseini

Rank	Author	Method	Q
1	HB	CVN Fourier	211
2	AL	KK99	131
3	TK	Lensfit	119
4	CH	KSBf90	52.3
5	PG	gfit	32.0
6	MV	KKshapelets with flexion	28.6
7	KK	KKshapelets	23.0
8	HHS3	GaussStackForwardGaussCleaned	22.4
9	SB	im2shape	20.1
10	HHS2	GaussStackForwardGauss	19.9
11	HHS1	Gauss	12.8
12	MJ	BJ02 deconvolved shapelets	9.80
13	USQM	USQM	1.22

Table 6.6: RealNoise_Blind leaderboard at the close of the challenge (Bridle *et al.* 2010*b*).

and M. Bethge, HB. This method was inspired by the second-place AL method, but improves on a key limitation which was highlighted by Lewis (2009) in that it did not depend on the galaxy centroid.

Fig. 6.7 shows Q as a function of galaxy type, PSF type, SNR, and galaxy size for RealNoise_Blind. The winning method, HB, performs well on all branches of the simulations but, interestingly, the second placed AL method performs better on six of the nine branches, including the fiducial simulation set. A poor performance on the low Signal-to-Noise branch heavily penalised the AL method, pushing it into second place.

The TK method does well when the bulge and disk galaxy profiles are separated, a result which stems from its use of co-aligned exponential and de Vaucouleurs components. The MV method does well in the high Signal-to-Noise regime.

Fig. 6.8 shows multiplicative and additive shear measurement bias as a function of galaxy type, PSF type, SNR, and galaxy size for RealNoise_Blind. The results broadly agree with previous conclusions from the STEP challenges. PSF type affects c strongly but not m. Incorrect estimation of PSF size tends to produce multiplicative bias but the GREAT08 PSFs were all roughly the same

size, varying only in ellipticity.

The HB method shows an overestimation of the true shear at about the 2% level. This is above the requirements for future surveys but seems to be approximately constant across all simulation branches. If a method could be shown to have a constant multiplicative bias, then it could be trivially removed from measurements of real data. The additive bias for HB is always below the requirement for far future surveys, $c < 0.0003$, except for one PSF branch and the low SNR branch. It will be important to chart the change in these values as the method is applied to more realistic simulations.

In general, the success of stacking methods is not a surprise, given the fact that GREAT08 `FITS` images provide a large number of well separated galaxies with constant shear applied. The future use of more realistic images may see a natural limit to the usefulness of stacking. However there are certain applications, including clusters of galaxies where stacking seems to be a natural analysis choice.

The fact that different methods submitted worked well in different areas of simulation parameter space offers hope for the prospect of future improvement in several of the techniques currently being explored. In practice a future practical shear estimation suite might combine the most successful elements from very different techniques. It is also encouraging that the winning entry was produced by a computer science student with no previous experience of astronomy, let alone weak gravitational lensing. This is something of a vindication of the GREAT08 concept and shows there certainly exists untapped experience and expertise relevant to cosmic shear beyond the WGL community.

6.2.6 GREAT10

The successor challenge to GREAT08 is currently underway (Kitching *et al.* 2010). It is called GREAT10 and seeks to build upon the success of the GREAT08 challenge while extending the problem to more sophisticated simulated images and other elements of the entire shear estimation pipeline. The orientation of the challenge towards the computer science community and other groups outside of traditional cosmology is still very firmly emphasised.

The major advancement in terms of simulated images is that shear is now allowed to vary across a single `FITS` image. This is much more like the data from real cosmic shear surveys. It will be every interesting to see the impact of variable shear on stacking techniques which rely on averaging large numbers of galaxies with identical shear.

There is a new component to the GREAT10 structure, complementary but separate from the

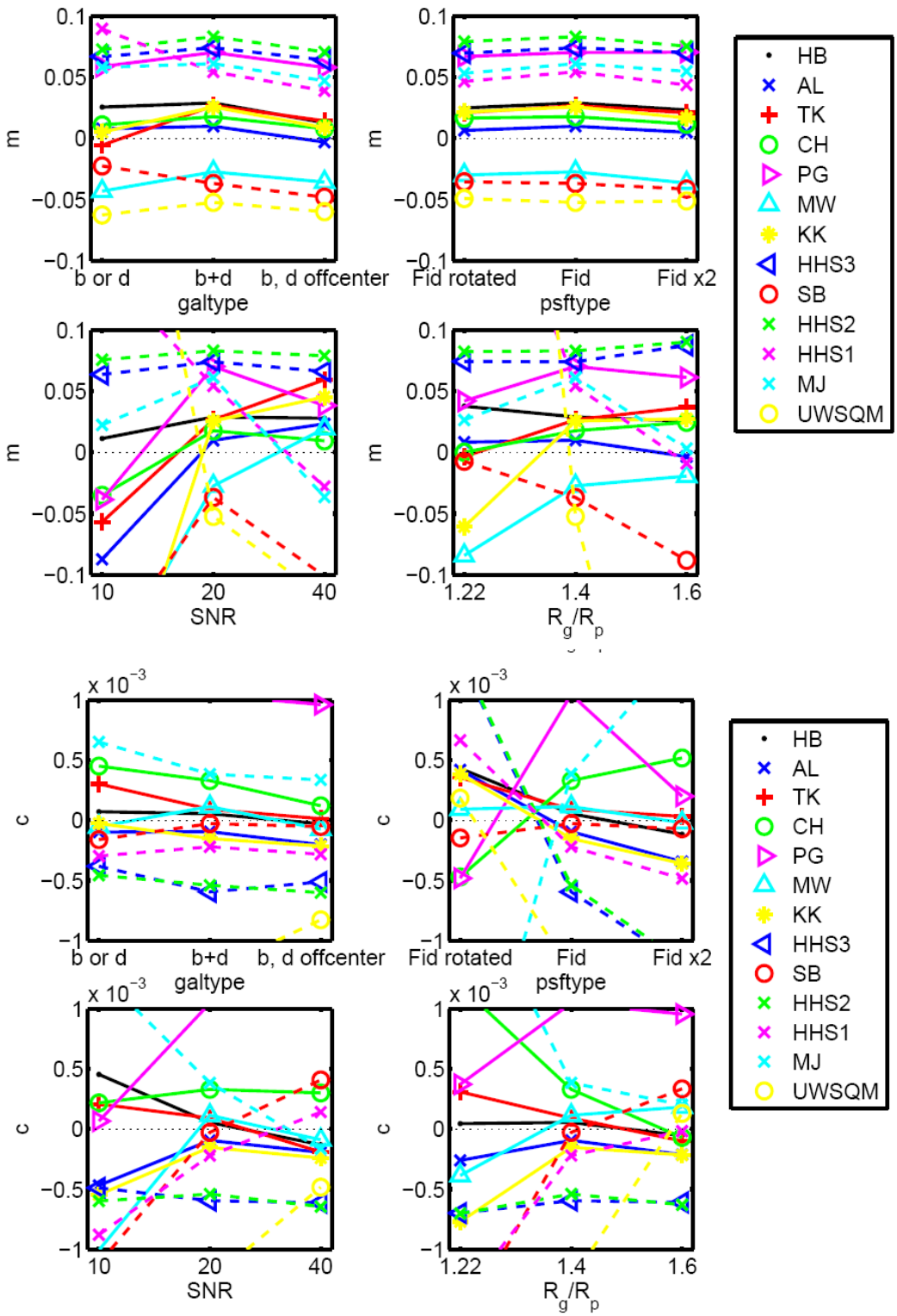

Figure 6.8: Upper Panel: Multiplicative shear measurement bias as a function of survey branch for RealNoise_Blind. Lower Panel: Additive shear measurement bias as a function of survey branch for RealNoise_Blind. (Bridle *et al.* 2010*b*)

galaxy challenge. This "star challenge" provides images of star fields which sparsely sample an underlying PSF. The PSF varies across the image field and the challenge is to reconstruct the PSF at non-star locations via interpolation.

The challenge is due to close in late 2011 with the results to be announced shortly after. The collaboration will then produced a detailed analysis of the competing methods and results, similar to that for GREAT08.

6.3 DES

The Dark Energy Survey will cover 5000deg^2 of the sky using a ground based telescope, mounted with a purpose built camera. The primary science goal of the survey is to better understand the nature of dark energy and the accelerating expansion of the Universe. To this end DES combines four probes of dark energy- Type Ia Supernovae; Baryon Acoustic Oscillations and Weak Gravitational Lensing. The survey is due to commence in late 2011 and continue for five years.

As a PhD student working in a DES institution I have been an active member of the Weak Lensing Working Group (WLWG). In particular, I have been responsible for evaluating the data storage and retrieval architecture of the collaboration and the evaluation of simulated data meant to test analysis pipelines in advance of the real survey data.

In 6.3.1 we briefly summarise the DES concept, collaboration and activities in preparation for the survey commencement. In 6.3.2 we discuss the tools produced by data management to allow storage and dissemination of DES simulated and survey data. We summarise development work and evaluation studies in the context of a large international collaboration. In 6.3.3 we detail the simulated images and catalogues of the DES Data Challenges (DCs) and present some of the results produced as part of the WLWG evaluation of DC5.

6.3.1 The Concept

The Dark Energy Survey aims to catalogue the sky in a 5000deg^2 area over 525 nights of viewing using the new Dark Energy Camera (DECam) mounted on the Blanco 4-meter telescope at the Cerro Tololo Inter-American Observatory, in the Chilean Andes. It will record information on about 300 million galaxies in five grizY optical passbands and reach g = 24.6, r = 24.1, i = 24.3, z = 23.8, and Y=21.5 AB at 10σ. The science goals of the survey primarily involve better understanding the nature of dark energy and the accelerating expansion of the Universe. DES science will consist of four probes- photometric supernovae, BAOs, galaxy clusters and WGL. The survey will be a leader in each of these disciplines and produce a large quantity of valuable science results beyond the constraints on DE.

The DES Collaboration is the organisation charged with successfully designing the DES project, building the DECam, conducting the survey and analysing the data. Comprising more than 120 collaborating scientists from 23 organizations, the collaboration is an example of the large and complex teams required to meet the demand for data in modern cosmology. The science work of the collaboration is organised through Science Working Groups for each of the four DES probes,

as well as groups for related topics, for example Combined Probes. DES Data Management are responsible for creating the data architecture to allow easy storage and retrieval of the terrabytes of data which the survey will produce as well as processing the data and producing catalogues of sources.

The DES observing strategy involves 80-100s exposures over 525 observing nights. The survey will cover $\sim$5000deg^2 in the southern sky. Multiple overlapping tilings of the same area will help optimise photometric redshift calibration. The DES observing area has been chosen to overlap with existing surveys including the South Pole Telescope (SPT), Sloan Digital Sky Survey (SDSS) and the Vista Hemisphere Survey (VHS). The overlap allows us to use information from complementary surveys to learn more about the objects DES will observe.

6.3.2 Evaluation of Data Management Tools

The Dark Energy Survey Data Management (DES:DM) team are responsible for developing the systems which will be used for the processing and calibration of DES data, and the DECam Community Pipeline, which will be used by the US National Optical Astronomy Observatory to process DECam data obtained by non-DES observers. The DESDM team will also operate the system during the survey. They are also responsible for co-ordinating the production of a number of increasingly complicated simulations of DES-type images and catalogues. These simulations are designed to test the analysis pipelines of the science working groups but they serve the additional purpose of testing the DES:DM data architecture- its ability to store data in the necessary volumes and to allow easy access to the data for members of the collaboration. This is complicated by the fact that different science working groups and members of the collaboration will have vastly different data needs. Some may require the raw nightly images for the whole survey, while others may only wish to access small patches of sky in the form of processed (cleaned, co-added) images, or even just object details in catalogue form.

To this end the DES:DM team created a number of tools to allow collaboration members to access data easily and in a convenient format. These included the *Archive Portal:* a central access point for all DES data products. The portal allows users to query the DES database and retrieve image files or catalogues for specified areas of the survey. *Control Portal:* allows direct grid transfer of large volumes of DES files via user-generated SQL queries. *Schema Browser:* provides information on the tables and variables included in the DES:DM database as well as details of how the tables are indexed.

As part of pre-survey testing the DES:DM tools were used to disseminate the series of DES

simulated images and catalogues known as Data Challenges (DCs). Each science working group was tasked with providing a number of people to test the working of the DES:DM tools by attempting to access the DC data and evaluate the results. I was asked to participate in this process as part of the WLWG, some conclusions are detailed below.

6.3.2.1 DES:DM Tools

The first DES:DM tool to be released to selected members of the collaboration was the DES Archive Portal. A screenshot is shown in Fig. 6.9. Initial evaluation was carried out during the period of Data Challenge 4 (DC4). The Portal interrogates the DC4 simulations database through two search pages, one for images and the other for objects i.e. catalogues. Each search allows the user to specify a range in ra/dec and options from a series of filters including Image ID, band, CCD no. etc. The search returns all objects matching the slected criteria. Individual objects (or the whole list) can then be slected for download. Different categories of images/catalogue are available: Objects, Co-Add Objects, Truth Objects, Shears, PSF Stars.

A number of shortcomings were identified from this evaluation. A general note was the lack of useful documentation. Much basic information was missing and the information that was provided tended to be in a very technical and obscure format. This was put right over several improvement cycles, leading eventually to the creation of the Schema Browser. Several code bugs which led to the return of incorrect or duplicated information were identified and rectified. On a structural level, the Archive Portal is a html-based online resource. Queries must be specified by hand and submitted one at a time and there was a limit of $\sim$4deg^2 on the total area that can be accessed in a single query. This is useful for teams including Cluster Lensing who require small numbers of small patches on the sky. It is much less convenient for collaboration members interested in acquiring either the whole simulated or surveyed area or in accessing a large number of smaller patches or individual object images.

On the recommendation of our assessment, a team was tasked with improving this aspect of the DES:DM system. The result was the creation of the DES:DM Control Portal. Initially this allowed the direct querying of the database through SQL queries and removed the maximum area limit. Further refinement allowed an automated query system via gridFTP which makes it significantly easier to access large numbers of individually specified images or area catalogues.

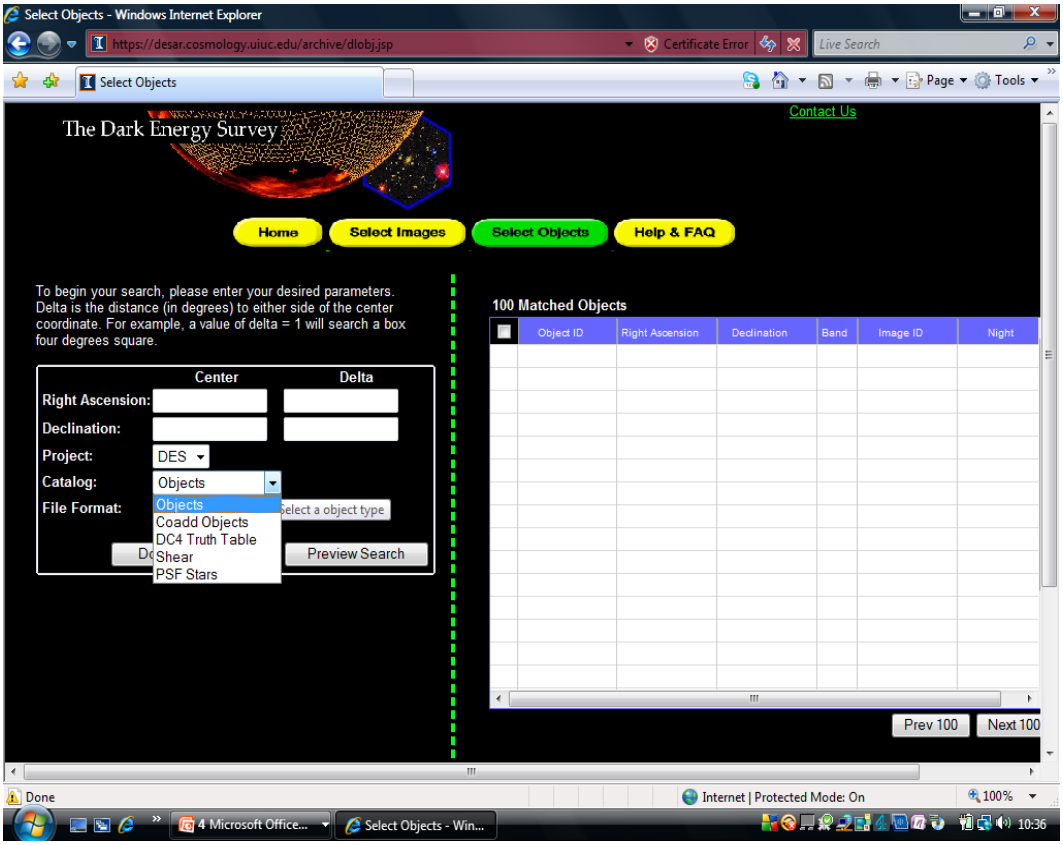

Figure 6.9: Screenshot of the DES:DM Archive Portal, http://cosmology.illinois.edu/archiveportal. This screen shows the Object Search form which will return catalogue information from a particular area of the survey. The active drop down menu allows the user to select the type of objects that will be returned in the catalogue.

6.3.2.2 The DES Brazil Portal

As well as evaluating the operation of the DES:DM tools, I also operated as the WLWG liaison with another data storage, processing and retrieval interface. The DES Brazil Portal, created and administered by the DES Brazil consortium, was conceived as a complementary structure to the DES:DM tools discussed above. The Brazil Portal is designed to link directly to the DES:DM database and allow users to access particular images or catalogues. Users can add code to their own workspace on the Brazil Portal and design a front-end user interface to allow others to access that code and run it using particular elements of the DES data- simulations in the first instance and eventually the DES survey data itself. There was some discussion that the DES Brazil Portal, if successful, could become the primary site for the science working groups to run their analyses of the final DES data products.

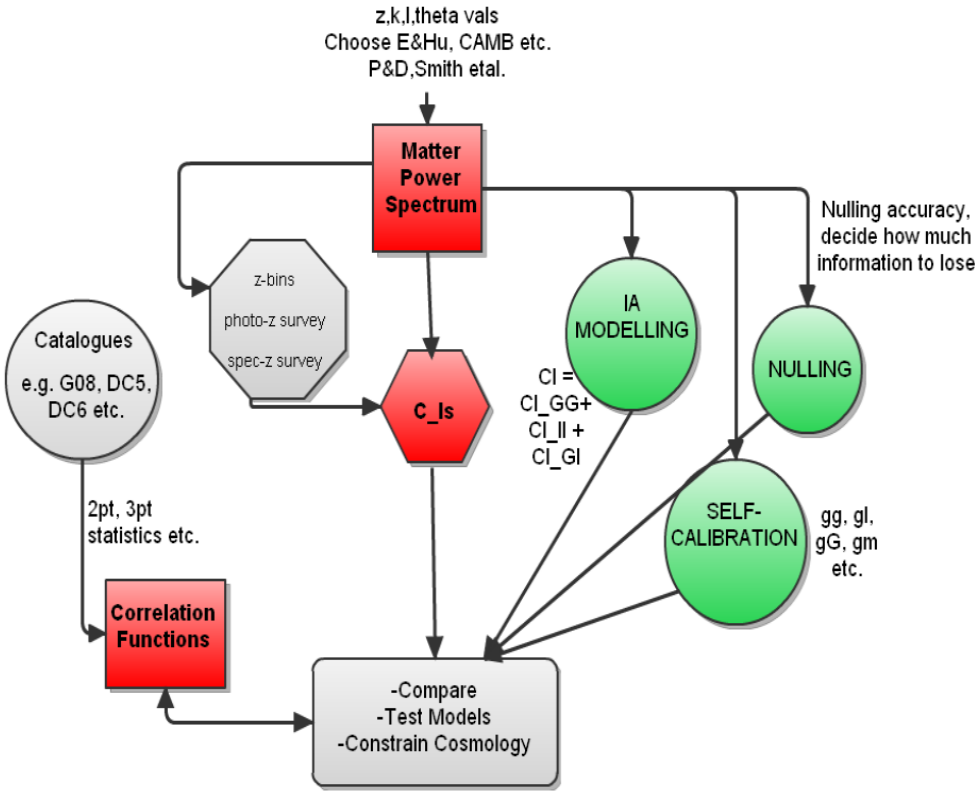

Figure 6.10: Proposed pipeline for integration of DES WLWG resources into the Brazil Portal Framework. Automated access ot the full DES:DM image and catalogue databases makes it realtively straightforward to run an automated, linked chain of code on particular data. An additional layer of complexity would allow for the treatment of systematic effects and the measurement of constraints on cosmological parameters.

As liaison between the WLWG and Brazil Portal we were asked to evaluate the operation of the portal and consider its potential uses towards the science goals of the WLWG. Fig. 6.10 shows a flow-chart detailing the different code components that the WLWG could add to the Brazil Portal. A full implementation of this scheme would allow a user to access a basic html interface involving drop-down selection boxes and a limited number of survey parameter options. After selecting a particular data source (DCs, full DES info, particularly survey nights) and area range, the user could calculate quantities of interest including correlation functions, power spectra, galaxy distributions etc. A more advanced implementation would allow the inclusion of IA contributions and the statistical comparison of fiducial cosmological models with simulations/data, perhaps even producing constraints on cosmological parameters. The code to allow the processes listed in the chart already exists across the DES collaboration, implementation in the Brazil Portal framework

would mainly involve the creation of a self-consistent user interface.

An unresolved question regarding this project centred on the practical benefits of conducting this work within the portal framework rather than running code privately and collaborating through the WLWG. The Portal front-end infrastructure makes it very straightforward for a relatively inexperienced user to run multiple complex analysis tools on any dataset within the DES:DM database. On the other hand, it may produce a "black box" mentality where users generate results without detailed knowledge of the workings and implicit assumptions of the pipeline code.

6.3.3 Data Challenges

As well as testing the functionality of the DES:DM tools, it was a main task of the DES WLWG to download the DES Data Challenge simulations and evaluate the contents. This acted both as a measure of the accuracy of our pipelines for shear analysis, which will eventually operate on the full DES dataset, and as a test of the DC simulations themselves through comparison of our results with their input catalogues. DES:DM have produced increasingly sophisticated suites of simulated images based on extensive N-body simulations. The improved simulation cycles are roughly annual. The latest iteration is DC6, the first to include both large-scale (200deg^2) and small-scale (20deg^2) datasets.

Our most comprehensive analysis was conducted on the DC5B simulated images, which did not include a large-field component. DC5B is the second release of the DC5 simulated data after some important bug-fixes. We accessed the catalogues of simulated objects via the DES:DM Archive portal, making $\sim$400 individual queries to keep the size of each data package manageable. At this stage there was still a relatively high error rate in the download process so many smaller sets mitigated the failure of individual downloads. The simulated galaxy images have been passed through a shear estimation pipeline (Jarvis) which uses a Principle Component Analysis to remove PSF effects and measure shear.

After download the information of interest, specifically shear but also star/galaxy flags and position information among other fields, could be extracted and the individual catalogue files stitched together to cover the full simulated area. After excluding those objects flagged as stars we can begin to analyse the galaxy shear information. The upper panel of Fig. 6.11 shows a shear-stick plot reconstructed from the DC5B galaxy information.Each line, or stick, on this plot represents a galaxy object, centred on the galaxy's RA/DEC position. The length of the stick is proportional to the magnitude of the shear signal for that galaxy ($|\gamma|^2 = \gamma_1^2 + \gamma_2^2$), the orientation gives the position angle of the shear. The DC5B filed spans a rectangular area of $\sim$ 16deg^2 with a

masked-out region of $\sim 0.75\text{deg}^2$.

The lower panel of Fig. 6.11 shows a simple histogram of the γ_1 component of the shear for the DC5B galaxies. The distribution is generally Gaussian around zero, this is consistent with the input to the simulated images. The most prominent feature, however, is a strong spike at $\gamma_1 = 0$. This is totally inconsistent with the input parameters and an unphysical result for a real galaxy sample. The shear measurement process may not be robust for some subset of the simulated galaxy shape distribution. If the measurement process fails catastrophically it could produce zero measured shear for some class of galaxy images. This unusual result and possible diagnoses were passed back to the DES:DM team for analysis and improvement during future Data Challenge iterations.

After discarding the galaxy entries with zero measured shear, we calculated the shear two point correlation function as an example of the statistical analysis the real shear data from DES will undergo as we try to measure cosmology. I wrote my own inefficient two point code and checked it against a fast 2d-tree correlation function code provided by Martin Kilbinger, a member of the WLWG, called athena [2]. On small areas ($< 0.1\text{deg}^2$) my code agreed with athena to within 20%, further agreement would not necessarily be expected due to the significant noise contribution from the small number of galaxies sampled. Applying my code to large areas proved prohibitive due to computer- and time-constraints. The agreement was thought sufficient to justify using athena for further DES data challenge analysis.

As expected the shear cross-term correlation function ($\xi_\times = \langle\gamma_\times\gamma_\times\rangle_\theta$) and the ξ_- correltation function ($\xi_- = \langle\gamma_t\gamma_t\rangle_\theta - \langle\gamma_r\gamma_r\rangle_\theta$) are both consistent with zero as we would expect. This is expected by symmetry: for example there is no reason why a pair of galaxies should consistently be pointing at the same positive 45 degree angle relative to the line joining them. This type of distortion cannot be created by matter in a single lens plane. Due to multiple deflections this can happen when there is more than one lens plane but the effect is very small. Also we expect the stretching along the line joining the galaxies to be the same size as the stretching perpendicular to this line, because the direction of the stretch is determined by the random position of the dark matter clump relative to the pair of galaxies. The ξ_+ correltation function ($\xi_+ = \langle\gamma_t\gamma_t\rangle_\theta + \langle\gamma_r\gamma_r\rangle_\theta$) is everywhere firmly non-zero and falls off towards larger scales. This result proved that the DC5B simulations could include a physically distributed shear field in the creation of galaxy images which was amenable to 2pt statistical analysis even over a small field on only $\sim$16deg^2. This was then compared to the input shear two point function and found to be significantly higher. The

[2] `http://www2.iap.fr/users/kilbinge/athena/`

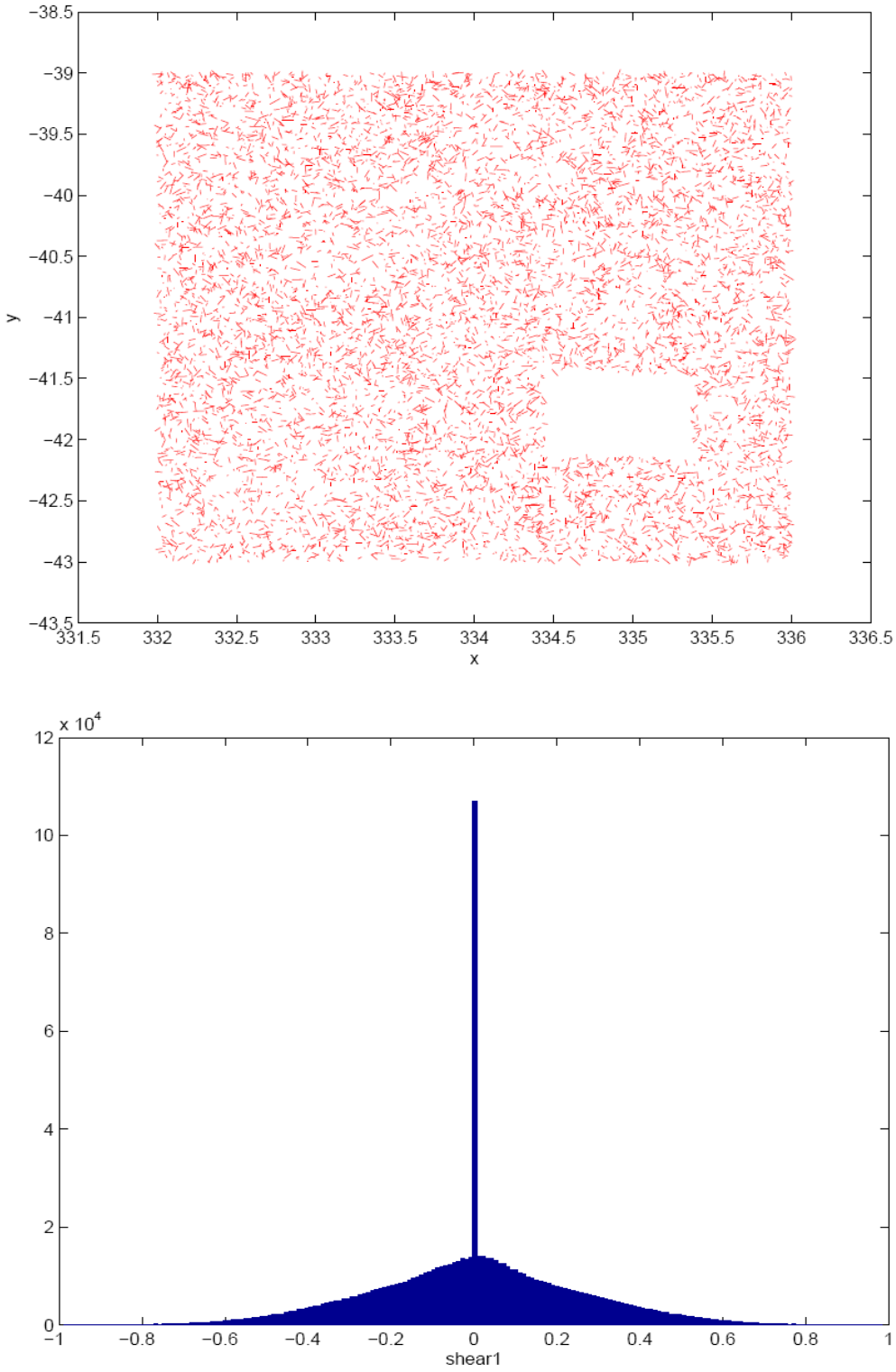

Figure 6.11: Upper Panel: Shearstick plot of all agalxy objects in the DC5B co-add catalogue. Stick position centred on object position, stick length proportional to absolute shear and orientation to the position angle of the shear relative to vertical. Empty patch has been masked out by the DES DM team due to faulty data processing. Lower Panel: Histogram of catalogue γ_1 values for all DC5B objects flagged as galaxies.

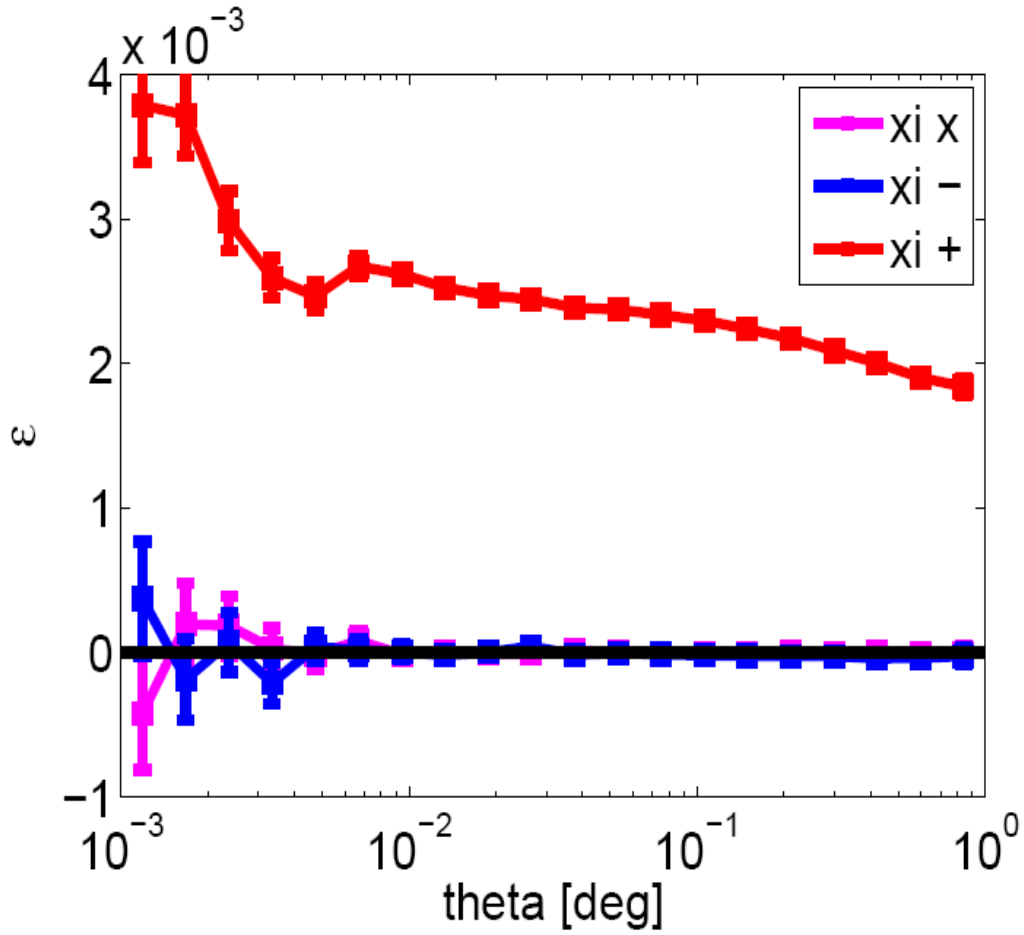

Figure 6.12: Shear-shear 2pt correlation functions for DC5B galaxies as a function of angular separation.The Red line is ξ_+, the sum of the radial and tangential correlation functions, The Blue line is ξ_-, the difference of the radial and tangential correlation functions and the Magenta line is $\xi_\times$, shear cross-term correlation function. Error bars give 68% confidence limits.

reason is still being investigated but a bug was found in the shear measurement pipeline shortly after the calculations were done, and it is now being re-run.

The Future The latest iteration of the DES Data Challenges, DC6, has now been released to the collaboration. As the first with a large area ($\sim$200deg^2) of processed shear measurements it will be an excellent test of our current shear measurement and anlysis pipelines. In the near future the first of a different set of simulations will be released to the collaboration. Dubbed the "Blind Cosmology Challenge", this dataset will have some assigned shear field but the shear values will be kept secret. The goal is for members of the collaboration to run their shear estimation codes on

the simulations and attempt to retrieve the shear information at the level of accuracy required by the DES science goals. The team at UCL hopes to take a central role in this task.

CHAPTER 7

CONCLUDING REMARKS & FUTURE WORK

This chapter aims to draw the themes of the thesis together in 7.1. It is natural that all research is, in some sense, open ended. Summarising the work contained in this thesis helps us to see directions in which the work could be extended. Section 7.2 highlights some of the possible future extensions including projects in related areas which could prove complementary to the direct future extensions of the work discussed.

7.1 Concluding Remarks

The goal of this work has been to explore the potential of cosmic shear to act as a *practical* probe of Modified Gravity (MG). A naive analysis of cosmic shear statistics puts relatively strong constraints on MG. However, it should never be acceptable to quote Weak Gravitational Lensing (WGL) measurements of cosmological parameters without accounting for the main astrophysical systematic contaminant of the cosmic shear signal, the Intrinsic Alignment (IA) of galaxy shapes. This thesis aims to demonstrate how one should best use cosmic shear data to test gravity, even with IAs included.

- In chapter 2 we illustrated the impact of IAs on current constraints from cosmic shear. We found that intrinsic alignments can either reduce or increase measurements of the fluctuation amplitude parameter σ_8 depending on the cosmic shear survey properties. The biasing of cosmological measurments due to cosmic shear was significant, on the order of $\sim 1 - 2\sigma$. We removed the IA biasing by marginalising a simple 2 parameter model, with the downside

of reduced constraining power. For the first time we here combine constraints from cosmic shear surveys (shear-shear correlations) with those from "GI" intrinsic alignment data sets (shear-position correlations). We find that the additional freedom due to marginalisation is well compensated by the additional information, in that the constraints are very similar indeed to those obtained when intrinsic alignments are ignored, both in terms of best fit values and uncertainties.

- We extended the IA formalism to include galaxy position-position nn and position-shear $n\epsilon$ correlations, reformulating the Linear Alignment model to be more physically consistent. The IA model is also expanded to contain more than 100 nuisance parameters, making it a much more reliable test of cosmology. We find that the constraining power of $\epsilon\epsilon$ alone is degraded by up to $\sim$75% by the inclusion of IAs. Including the galaxy position information goes some way towards mitigating the IA effect, doubling the figure of merit (FoM) between $\epsilon\epsilon$ and $\epsilon\epsilon + n\epsilon + nn$. We replicate these results for constraints on two parameters, Q_0 and R_0 which quantify deviations from General Relativity. We present the first forecast constraints on Modified Gravity (MG) from cosmic shear including IAs, showing that the behaviour is very similar to that of the dark energy constraints. We find that these results are robust over a range of bias parameters.

- To constrain dark energy and MG in practice from cosmic shear, we rely on large WGL surveys. Chapter 4 explores the impact of survey configuration and redshift information on our ability to constrain cosmology in the presence of IAs. We find that increased area improves the FoM. However, when we model a realistic situation where fixed total survey time means that increased area demands a lower density of source galaxies on the sky, we see a very different effect. Now there is little improvement from increasing area above $\sim$5,000deg^2. In fact, MG constraints get worse at larger areas. We attribute this to the fact that the survey must be shallower to keep a fixed survey time, and IAs dominate at low redshift. This preference for smaller areas is contrary to the result for dark energy constraints without IAs.

  The impact of improved redshift accuracy is far from dramatic due to the fact that the lensing integral smears out redshift information. MG constraints benefit somewhat more than those for DE, suggesting that greater resolution in redshift helps to constrain the fixed z-variation

of our MG model.

- Beyond the possibilities for "self-calibration", improving constraints with data already implicit in cosmic shear surveys, there is great scope for the production of joint constraints from multiple complementary probes. As the number of probes and amount of data increases over the coming decade, the production of correctly calculated joint constraints will become increasingly important. Joint constraints allow parameter degeneracies to be broken and have the potential to produce constraints which are orders of magnitude tighter than any single probe alone. We demonstrate that the addition of a purely spectroscopic survey to improve the redshift accuracy of a WGL photometric survey has negligible impact compared to the cost of measuring millions of galaxy spectra.

  However, when we use a smaller spectroscopic survey to measure the Redshift Space Distortions (RSDs) in the galaxy power spectrum, we see that the joint WGL-RSD constraint on dark energy is several times better than the RSD constraint alone. The results are even more striking when we attempt to constrain MG. WGL responds to the sum of the metric potentials, $\Phi + \Psi$, while RSDs are sourced by the Newtonian potential, Ψ, alone. The combination of the probes allows degeneracies to be broken and produces constraints on MG which are orders of magnitude stronger than either probe alone. We hope to refine these results in the context of DESpec by accounting for the cross-correlations due to both probes being measured on the same patch of the sky.

- Chapter 6 summarises some work contributed to two international collaborations. Dark Energy Survey (DES), a stage-III cosmology survey due to see first light in late 2011 is an example of the large collaborations that now form the backbone of modern cosmology. Combining four probes: WGL; Supernovae; BAOs and Clusters, DES will produce some of the tightest constraints on cosmology. We discuss some of the work contributed to the planning of this mission. We summarise our evaluation of the extensive DES: Data Management architecture for storing and disseminating DES simulated and, eventually, real survey data. We offer constructive criticisms of a novel survey data interface, the DES: Brazil portal and present the results of our shear 2pt statistics analysis of a set of simulated DES images, Data Challenge 5.

7.2 Future Work

This section outlines possible extensions to the work considered in this thesis. Our results have shown that weak gravitational lensing, in combination with galaxy position information, is a powerful probe of gravity, even in the presence of large systematic effects. As new data becomes available over the next decade there is an excellent opportunity to expand this work and create a coherent, comprehensive framework for the production of optimal cosmological constraints from WGL.

7.2.1 Intrinsic Alignments

In this thesis we began with a very minimal intrinsic alignment model of just two parameters and moved to a joint IA/galaxy biasing model with a fiducial number of 104 nuisance parameters. This vast increase aims to ensure that we do not over-constrain our knowledge of IAs. That we can still produce scientifically useful results with such a broad bias model, able to vary in scale and redshift, is encouraging.

In the future we would like to use less brute force when parameterising IAs. An improved knowledge of the physics of IAs, especially at small scales, is one avenue of interest. We have gone some way in this direction in chapter 2 by building on the work of Schneider & Bridle (2010*a*) who use the halo model of galaxy clustering to model the non-linear IA of galaxies. Extending this work over a range of fiducial cosmologies is one basic step that could assist with modelling of IAs. More ambitiously, theoretical and observational advances in the study of galaxy formation may help us learn about the IA effect. Particularly important is to study galaxy IA as a function of galaxy properties - galaxy type, luminosity etc, and match these to the composition of particular surveys.

It is possible that a better understanding of IAs may let us reduce the number of nuisance parameters in our IA model, or at least place some informative priors on them. This would help reduce their impact on the constraining power of cosmic shear surveys.

7.2.2 Comprehensive Self-Calibration Programme

The use of galaxy position data, already present in any cosmic shear survey, to mitigate the impact of IAs is an example of "self-calibration" of a cosmic probe. There is a developing programme in the WGL community to develop self-calibration techniques. We have drawn extensively on the work of (Joachimi & Bridle 2009). We would aim to include the magnification effects of IAs

into our framework. The work of Zhang (2010); Zhang *et al.* (2010) has suggested that WGL self-calibration could extend to improvements on photo-z accuracy and the use of high-resolution tomographic analysis to remove IA effects.

The self-calibration idea has generated a lot of interest and some ingenious plans. There is a danger that much piecemeal work could end up "double-counting" the information in cosmic shear surveys, producing over-optimistic constraints if we naively believed we could implement all techniques simultaneously. We aim to integrate the suggested self-calibrations into a single, coherent and self-consistent formalism which aims to optimise the constraining power of a cosmic shear survey.

The combination of WGL at different redshifts is a possible useful extension of this idea. Brown & Battye (2011*b*) have suggested that polarisation measurements from radio WGL surveys could provide information on the intrinsic ellipticity of source galaxies. If sufficient source galaxies could be detected this has the potential to strongly constrain IA effects and improve constraints from cosmic shear.

7.2.3 Modified Gravity

In contrast to the IA model, our parameterisation of MG was relatively simple. The simple Q_0, R_0 parameterisation is an excellent test of concept. If we were able to constrain either parameter away from their GR value of 1 (in the absence of systematics), it would indeed be a momentous discovery. In future though we would hope to extend the model, making it more flexible. This will have the double benefit of allowing us to capture more information about any deviation from GR and, in the case of a detection of beyond-GR behaviour, may provide more clues towards a new theory of gravity. The inevitable downside is that more free parameters will produce looser constraints.

One possibility is to allow a freedom in both scale and redshift space, as per our IA model. It may be possible to apply a principle component analysis to identify the best way to parameterise modes in k/z space. Alternatively parameterisations linked to specific theories of MG could be included. In this way theoretical priors could cut down the large parameter space we would need to explore.

Alternatively it would be interesting to investigate the meaning of fitting our simple paramaterisation to the data even in the case where the underlying model is more complicated. Simpson & Bridle (2005, 2006) demonstrated that constraining only one or two dark energy parameters measured a particular weighted average of the underlying dark energy equation of state as a function

of redshift, to a good approximation. The same may be true of a fitted Q_0 and R_0 value.

In this thesis we have assumed that any MG theory of interest would exactly mimic the observed expansion history of the Universe, behaving like an effective dark energy. The ability to discriminate between dark energy and MG comes from the different predictions for the growth of structure. The behaviour at non-linear scales is at present rather ill-understood but could be of considerable value. We have assumed that the Smith et al. non-linear fitting function work in the MG regime. Other groups have calculated a particular fitting function for some value of MG parameters and interpolated with the Smith et al. result for a ΛCDM Universe. Future efforts towards modelling non-linear clustering behaviour in the presence of some generic deviations from GR could be an extremely beneficial addition to our cosmological toolbox.

7.2.4 DES, DESpec & Future Surveys

The work of this thesis has demonstrated not only the potential for the practical use of cosmic shear as a cosmological tool in combination with other information implicit in WGL surveys, but also the complementarity between cosmic shear constraints and other probes, principally Redshift Space Distortions. The DES/DESpec synergy was striking but the picture is, as yet, only half complete.

The cosmic shear and RSD measurements were added as if they were independent. This misses one of the main points of the DESpec proposal - that it covers the same patch of sky and shares its source catalogue with a large photometric survey. There are sure to be causal connections between the physics responsible for the galaxy clustering and peculiar velocities at low redshifts and the WGL effect on sources at higher redshift. These connections give rise to cross-correlations in the data. To treat these fully we must drop the idea that the probes are independent and calculate the full covariance matrix, including off diagonal elements.

To do this it will be easiest if we re-cast the RSD components in C_l language using spherical harmonics, taking care to fully include the Kaiser effect. This will enable us to use the covariance matrix formalism found in chapter 3.

This fully worked out cross-correlation approach is easily generalisable to other probes. Weak gravitational lensing of the CMB is an obvious example of a correlation between early Universe physics and late time large scale structure. Cross correlations with the ISW effect could also be fruitfully probed.

The over-arching goal, as with the self-calibration programme, should be to strive for a coherent and comprehensive formalism for the cross-correlation of cosmic probes. It is these accurate

joint constraints that will not only produce the tightest constraints for future surveys but also reveal tensions in models and the existence of systematics through the measurement of "off-centred" best-fit cosmologies.

We have explored the impact of spectroscopic redshifts on WGL in some detail. This can, of course, be applied to other cosmic probes. Our analysis was useful as a first order discussion of the relative effect of spectroscopic redshift information on cosmic shear and related probes. However, to fairly assess the impact we should optimise the employment of accurate redshift information, re-analysing with an increased number of tomographic bins. It may also be the case that there is some optimal trade-off between using spec-z information to calibrate a particular photo-z distribution, and probe physics directly through e.g. RSDs.

We have demonstrated the usefulness of multiple probes to constrain deviations from GR. A different topic that could benefit from a similar treatment is constraints in the mass of the neutrino through observations of large-scale structure. Cosmology has the potential to put even more stringent constraints on the neutrino mass than particle physics lab-based experiments. The use of cosmic shear and galaxy position data together can help to break the degeneracy between neutrino mass and galaxy biasing - this project seems well suited to our lensing + galaxy position formalism.

7.2.5 Shear Estimation

No ambitious plans for cosmic shear, whether the modelling of IAs or self-calibration, will reach their full potential without improvements in our ability to estimate shear from galaxy images. While the GREAT08 challenge produced improvements on the best methods from within the WGL community, the final results were far from the $Q = 1000$ goal for future stage-IV type surveys.

The GREAT10 challenge asks participants to surpass previous efforts while tackling a more realistic, and complicated, set of simulated images. The success, or otherwise, of stacking-based methods will be very interesting. If they fail to deal with a shear field that varies across the image then we will be left with the family of techniques which fit the shape of individual galaxies. If the stacking methods show some resilience there will be a lot of interesting work to evaluate their potential success on real survey images, particularly the requirements they might place on projected number density and target selection.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A. & et al., 2009, *ApJS*, **182**, 543

Adelberger, E. G., Heckel, B. R. & Nelson, A. E., 2003, *Annual Review of Nuclear and Particle Science*, **53**, 77

Albrecht, A., Bernstein, G., Cahn, R., Freedman, W. L., Hewitt, J., Hu, W., Huth, J., Kamionkowski, M., Kolb, E. W., Knox, L., Mather, J. C., Staggs, S. & Suntzeff, N. B., 2006, *ArXiv Astrophysics e-prints*

Alcock, C., Allsman, R. A., Alves, D. R., Axelrod, T. S., Becker, A. C., Bennett, D. P., Cook, K. H., Dalal, N., Drake, A. J., Freeman, K. C., Geha, M., Griest, K., Lehner, M. J., Marshall, S. L., Minniti, D., Nelson, C. A., Peterson, B. A., Popowski, P., Pratt, M. R., Quinn, P. J., Stubbs, C. W., Sutherland, W., Tomaney, A. B., Vandehei, T. & Welch, D., 2000, *ApJ*, **542**, 281

Amara, A. & Refregier, A., 2006, *ArXiv Astrophysics e-prints*

Amara, A. & Refregier, A., 2007

Aubert, D., Pichon, C. & Colombi, S., 2004, *MNRAS*, **352**, 376

Aubourg, E., Bareyre, P., Bréhin, S., Gros, M., Lachièze-Rey, M., Laurent, B., Lesquoy, E., Magneville, C., Milsztajn, A., Moscoso, L., Queinnec, F., Rich, J., Spiro, M., Vigroux, L., Zylberajch, S., Ansari, R., Cavalier, F., Moniez, M., Beaulieu, J., Ferlet, R., Grison, P., Vidal-Madjar, A., Guibert, J., Moreau, O., Tajahmady, F., Maurice, E., Prévôt, L. & Gry, C., 1993, *Nature*, **365**, 623

Bacon, D. J., Refregier, A. R. & Ellis, R. S., 2000, *MNRAS*, **318**, 625

Bacon, D. J. & Schäfer, B. M., 2009, *MNRAS*, **396**, 2167

Bartelmann, M. & Schneider, P., 2001, *Phys. Rept.*, **340**, 291

Bean, R. & Tangmatitham, M., 2010, *PRD*, **81**(8), 083534

Benjamin, J., Heymans, C., Semboloni, E., Van Waerbeke, L., Hoekstra, H., Erben, T., Gladders, M. D., Hetterscheidt, M., Mellier, Y. & Yee, H. K. C., 2007, *MNRAS*, **381**, 702

Bernstein, G. M., 2009, *ApJ*, **695**, 652

Bernstein, G. M., 2010, *MNRAS*, **406**, 2793

Blake, C. & Bridle, S., 2005, *MNRAS*, **363**, 1329

Blake, C. A., Abdalla, F. B., Bridle, S. L. & Rawlings, S., 2004, *New Astron.Rev.*, **48**, 1063

Blandford, R. D., Saust, A. B., Brainerd, T. G. & Villumsen, J. V., 1991, *MNRAS*, **251**, 600

Bolzonella, M., Miralles, J. & Pelló, R., 2000, *A&A*, **363**, 476

Bond, J. R., Efstathiou, G. & Silk, J., 1980, *Physical Review Letters*, **45**, 1980

Bridle, S., Balan, S. T., Bethge, M., Gentile, M., Harmeling, S., Heymans, C., Hirsch, M., Hosseini, R., Jarvis, M., Kirk, D., Kitching, T., Kuijken, K., Lewis, A., Paulin-Henriksson, S., Schölkopf, B., Velander, M., Voigt, L., Witherick, D., Amara, A., Bernstein, G., Courbin, F., Gill, M., Heavens, A., Mandelbaum, R., Massey, R., Moghaddam, B., Rassat, A., Réfrégier, A., Rhodes, J., Schrabback, T., Shawe-Taylor, J., Shmakova, M., van Waerbeke, L. & Wittman, D., 2010*a*, *MNRAS*, **405**, 2044

Bridle, S., Balan, S. T., Bethge, M., Gentile, M., Harmeling, S., Heymans, C., Hirsch, M., Hosseini, R., Jarvis, M., Kirk, D., Kitching, T., Kuijken, K., Lewis, A., Paulin-Henriksson, S., Schölkopf, B., Velander, M., Voigt, L., Witherick, D., Amara, A., Bernstein, G., Courbin, F., Gill, M., Heavens, A., Mandelbaum, R., Massey, R., Moghaddam, B., Rassat, A., Réfrégier, A., Rhodes, J., Schrabback, T., Shawe-Taylor, J., Shmakova, M., van Waerbeke, L. & Wittman, D., 2010*b*, *MNRAS*, **405**, 2044

Bridle, S. & King, L., 2007, *New Journal of Physics*, **9**, 444

Bridle, S., Shawe-Taylor, J., Amara, A., Applegate, D., Balan, S. T., Berge, J., Bernstein, G., Dahle, H., Erben, T., Gill, M., Heavens, A., Heymans, C., High, W., Hoekstra, H., Jarvis, M., Kirk, D., Kitching, T., Kneib, J.-P., Kuijken, K., Lagattuta, D., Mandelbaum, R., Massey,

R., Mellier, Y., Moghaddam, B., Moudden, Y., Nakajima, R., Paulin-Henriksson, S., Pires, S., Rassat, A., Refregier, A., Rhodes, J., Schrabback, T., Semboloni, E., Shmakova, M., van Waerbeke, L., Witherick, D., Voigt, L. & Wittman, D., 2009, *Annals of Applied Statistics*, **3**, 6

Brown, M. L. & Battye, R. A., 2011*a*, *ArXiv e-prints*

Brown, M. L. & Battye, R. A., 2011*b*, *MNRAS*, **410**, 2057

Brown, M. L., Taylor, A. N., Hambly, N. C. & Dye, S., 2002, *MNRAS*, **333**, 501

Carroll, S. M., 2001, in J. A. Harvey, S. Kachru, & E. Silverstein (ed.), *Strings, Branes and Gravity TASI 99*, pp. 437–492

Carroll, S. M., Press, W. H. & Turner, E. L., 1992, *ArA&A*, **30**, 499

Catelan, P., Kamionkowski, M. & Blandford, R. D., 2001, *MNRAS*, **320**, L7

Chang, T., Refregier, A. & Helfand, D. J., 2004, *ApJ*, **617**, 794

Clowe, D., Bradač, M., Gonzalez, A. H., Markevitch, M., Randall, S. W., Jones, C. & Zaritsky, D., 2006, *ApJL*, **648**, L109

Collister, A. A. & Lahav, O., 2004, *Publ. Astron. Soc. Pac.*, **16**, 345

Cooray, A., 2006, *ArXiv Astrophysics e-prints*

Cooray, A., Hu, W. & Miralda-Escudé, J., 2000, *ApJL*, **535**, L9

Cooray, A. & Sheth, R., 2002, *Physics Reports*, **372**, 1

Crittenden, R. G., Natarajan, P., Pen, U.-L. & Theuns, T., 2001, *ApJ*, **559**, 552

Dodelson, S., 2003, *Modern Cosmology* (Academic Press)

Dunkley, J., Komatsu, E., Nolta, M. R., Spergel, D. N., Larson, D., Hinshaw, G., Page, L., Bennett, C. L., Gold, B., Jarosik, N., Weiland, J. L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E. & Wright, E. L., 2009, *ApJS*, **180**, 306

Durrer, R. & Maartens, R., 2008, *General Relativity and Gravitation*, **40**, 301

Dyson, F. W., Eddington, A. S. & Davidson, C., 1920, *Phil. trans. Royal soc.*, **220**

Einstein, A., 1916, *AnP*, **354**, 769

Eisenstein, D. J. & Hu, W., 1998, *ApJ*, **496**, 605

Feng, J. L., 2010, *ArA&A*, **48**, 495

Fu, L. *et al.*, 2008, *A&A*, **479**, 9

Glazebrook, K. & Blake, C., 2005, *ApJ*, **631**, 1

Grossman, S. A. & Narayan, R., 1988, *ApJL*, **324**, L37

Guth, A. H., 1981, *PRD*, **23**, 347

Guzik, J., Jain, B. & Takada, M., 2010, *PRD*, **81**(2), 023503

Hawkins, E., Maddox, S., Cole, S., Lahav, O., Madgwick, D. S., Norberg, P., Peacock, J. A., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Jones, B., Lewis, I., Lumsden, S., Percival, W., Peterson, B. A., Sutherland, W. & Taylor, K., 2003, *MNRAS*, **346**, 78

Heavens, A., 2003, *MNRAS*, **343**, 1327

Heavens, A. & Peacock, J., 1988, *MNRAS*, **232**, 339

Hetterscheidt, M., Simon, P., Schirmer, M., Hildebrandt, H., Schrabback, T., Erben, T. & Schneider, P., 2007, *A&A*, **468**, 859

Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F. & Langston, G. I., 1988, *Nature*, **333**, 537

Heymans, C. *et al.*, 2006, *Mon. Not. Roy. Astron. Soc.*, **368**, 1323

Heymans, C., Brown, M., Heavens, A., Meisenheimer, K., Taylor, A. & Wolf, C., 2004*a*, *MNRAS*, **347**, 895

Heymans, C., Brown, M., Heavens, A., Meisenheimer, K., Taylor, A. & Wolf, C., 2004*b*, *MNRAS*, **347**, 895

Heymans, C. & Heavens, A., 2003, *MNRAS*, **339**, 711

Heymans, C., Van Waerbeke, L., Bacon, D., Berge, J., Bernstein, G., Bertin, E., Bridle, S., Brown, M. L., Clowe, D., Dahle, H., Erben, T., Gray, M., Hetterscheidt, M., Hoekstra, H., Hudelot, P.,

Jarvis, M., Kuijken, K., Margoniner, V., Massey, R., Mellier, Y., Nakajima, R., Refregier, A., Rhodes, J., Schrabback, T. & Wittman, D., 2006, *MNRAS*, **368**, 1323

Hildebrandt, H. *et al.*, 2010, *A&A*, **523**, A31+

Hirata, C. M., Mandelbaum, R., Ishak, M., Seljak, U., Nichol, R., Pimbblet, K. A., Ross, N. P. & Wake, D., 2007, *MNRAS*, **381**, 1197

Hirata, C. M., Mandelbaum, R., Seljak, U., Guzik, J., Padmanabhan, N., Blake, C., Brinkmann, J., Budávari, T., Connolly, A., Csabai, I., Scranton, R. & Szalay, A. S., 2004, *MNRAS*, **353**, 529

Hirata, C. M. & Seljak, U., 2004, *PRD*, **70**(6), 063526

Hoekstra, H. *et al.*, 2006, *Astrophys. J.*, **647**, 116

Hoekstra, H., Yee, H. K. C., Gladders, M. D., Barrientos, L. F., Hall, P. B. & Infante, L., 2002, *ApJ*, **572**, 55

Hogg, D. W., 1999, *ArXiv Astrophysics e-prints*

Hu, W. & Jain, B., 2004, *PRD*, **70**(4), 043009

Hubble, E., 1929, *Proc. N. A. S.*, **15**, 168

Huterer, D., Takada, M., Bernstein, G. & Jain, B., 2006, *Mon. Not. Roy. Astron. Soc.*, **366**, 101

I. Lazslo, D. Kirk, R. Bean and S. Bridle, 2011

Ilbert, O. *et al.*, 2006, *A&A*, **457**, 841

Jaffe, A., 1996, *ApJ*, **471**, 24

Jain, B. & Khoury, J., 2010, *Annals of Physics*, **325**, 1479

Jing, Y. P., 2002, *MNRAS*, **335**, L89

Joachimi, B. & Bridle, S. L., 2009, *ArXiv e-prints*

Joachimi, B. & Bridle, S. L., 2010, *A&A*, **523**, A1+

Joachimi, B., Mandelbaum, R., Abdalla, F. B. & Bridle, S. L., 2011, *A&A*, **527**, A26+

Joachimi, B. & Schneider, P., 2008, *A&A*, **488**, 829

Joachimi, B. & Schneider, P., 2009, *A&A*, **507**, 105

Kaiser, N., 1988, *MNRAS*, **231**, 149

Kaiser, N., 1992, *ApJ*, **388**, 272

Kaiser, N., Squires, G. & Broadhurst, T., 1995, *ApJ*, **449**, 460

Kaiser, N., Wilson, G. & Luppino, G. A., 2000, *ArXiv Astrophysics e-prints*

Kaiser, N., Wilson, G. & Luppino, G. A., 2000

King, L. J. & Schneider, P., 2003, *A&A*, **398**, 23

Kirk, D., Bridle, S. & Schneider, M., 2010, *MNRAS*, **408**, 1502

Kitching, T., Amara, A., Gill, M., Harmeling, S., Heymans, C., Massey, R., Rowe, B., Schrabback, T., Voigt, L., Balan, S., Bernstein, G., Bethge, M., Bridle, S., Courbin, F., Gentile, M., Heavens, A., Hirsch, M., Hosseini, R., Kiessling, A., Kirk, D., Kuijken, K., Mandelbaum, R., Moghaddam, B., Nurbaeva, G., Paulin-Henriksson, S., Rassat, A., Rhodes, J., Schölkopf, B., Shawe-Taylor, J., Shmakova, M., Taylor, A., Velander, M., van Waerbeke, L., Witherick, D. & Wittman, D., 2010, *ArXiv e-prints*

Kneib, J., Ellis, R. S., Santos, M. R. & Richard, J., 2004, *ApJ*, **607**, 697

Kneib, J. P., Ellis, R. S., Smail, I., Couch, W. J. & Sharples, R. M., 1996, *ApJ*, **471**, 643

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E. & Wright, E. L., 2011, *ApJS*, **192**, 18

Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S. & Moustakas, L. A., 2006, *ApJ*, **649**, 599

Kristian, J. & Sachs, R. K., 1966, *ApJ*, **143**, 379

Lahav, O., 2001, *Organizations and Strategies in Astronomy, Vol. II*, **266**, 139

Lahav, O., Bridle, S. L., Percival, W. J., Peacock, J. A., Efstathiou, G., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., de Propris, R., Driver, S. P., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lewis, I.,

Lumsden, S., Maddox, S., Madgwick, D. S., Moody, S., Norberg, P., Peterson, B. A., Sutherland, W. & Taylor, K., 2002, *MNRAS*, **333**, 961

Landy, S. D. & Szalay, A. S., 1993, *ApJ*, **412**, 64

Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E. & Wright, E. L., 2011, *ApJS*, **192**, 16

Le Fèvre, O., Mellier, Y., McCracken, H. J., Foucaud, S., Gwyn, S., Radovich, M., Dantel-Fort, M., Bertin, E., Moreau, C., Cuillandre, J., Pierre, M., Le Brun, V., Mazure, A. & Tresse, L., 2004, *A&A*, **417**, 839

Lee, M. H. & Paczynski, B., 1990, *ApJ*, **357**, 32

Lewis, A., 2009, *MNRAS*, **398**, 471

Lewis, A. & Bridle, S., 2002, *PRD*, **66**(10), 103511

Lewis, A. & Challinor, A., 2006, *Physics Reports*, **429**, 1

Liddle, A. R. & Lyth, D. H., 2000 (Cambridge University Press)

Mandelbaum, R. *et al.*, 2009, *ArXiv e-prints*

Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U. & Brinkmann, J., 2006*a*, *MNRAS*, **370**, 1008

Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U. & Brinkmann, J., 2006*b*, *MNRAS*, **370**, 1008

Mandelbaum, R., Hirata, C. M., Ishak, M., Seljak, U. & Brinkmann, J., 2006*c*, *MNRAS*, **367**, 611

Mandelbaum, R., Hirata, C. M., Seljak, U., Guzik, J., Padmanabhan, N., Blake, C., Blanton, M. R., Lupton, R. & Brinkmann, J., 2005, *MNRAS*, **361**, 1287

Massey, R. *et al.*, 2007, *ApJS*, **172**, 239

Massey, R. *et al.*, 2007*a*, *Nature*, **445**, 286

Massey, R. *et al.*, 2007*b*, *Mon. Not. Roy. Astron. Soc.*, **376**, 13

Massey, R., Heymans, C., Bergé, J., Bernstein, G., Bridle, S., Clowe, D., Dahle, H., Ellis, R., Erben, T., Hetterscheidt, M., High, F. W., Hirata, C., Hoekstra, H., Hudelot, P., Jarvis, M., Johnston, D., Kuijken, K., Margoniner, V., Mandelbaum, R., Mellier, Y., Nakajima, R., Paulin-Henriksson, S., Peeples, M., Roat, C., Refregier, A., Rhodes, J., Schrabback, T., Schirmer, M., Seljak, U., Semboloni, E. & van Waerbeke, L., 2007, *MNRAS*, **376**, 13

Massey, R., Stoughton, C., Leauthaud, A., Rhodes, J., Koekemoer, A., Ellis, R. & Shaghoulian, E., 2010, *MNRAS*, **401**, 371

Mather, J. C., Fixsen, D. J., Shafer, R. A., Mosier, C. & Wilkinson, D. T., 1999, *ApJ*, **512**, 511

Metcalf, R. B., 2007, *Nuovo Cimento B Serie*, **122**, 1453

Milgrom, M., 1994, *Annals of Physics*, **229**, 384

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F. & van Waerbeke, L., 2007, *MNRAS*, **382**, 315

Miralda-Escude, J., 1991, *ApJ*, **370**, 1

Mukhanov, V. F., Feldman, H. A. & Brandenberger, R. H., 1992, *Physics Reports*, **215**, 203

Munshi, D., Kitching, T., Heavens, A. & Coles, P., 2011, *MNRAS*, 1248

Munshi, D., Valageas, P., Van Waerbeke, L. & Heavens, A., 2006

Narayan, R. & Bartelmann, M., 1996, *preprint (astro-ph/9606001)*

Pan, J. & Zhang, P., 2010, *ApJ*, **8**, 19

Paraficz, D. & Hjorth, J., 2010, *ApJ*, **712**, 1378

Patel, P., Bacon, D. J., Beswick, R. J., Muxlow, T. W. B. & Hoyle, B., 2010, *MNRAS*, **401**, 2572

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A. & Bridle, S. L., 2008, *A&A*, **484**, 67

Peacock, J. & Schneider, P., 2006, *The Messenger*, **125**, 48

Peacock, J. A. *et al.*, 2006

Peebles, P. J. E., 1993 (Princeton Series in Physics, Princeton, NJ: Princeton University Press, —c1993)

Peng, C. Y., Ho, L. C., Impey, C. D. & Rix, H., 2002, *AJ*, **124**, 266

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J. & The Supernova Cosmology Project, 1999, *ApJ*, **517**, 565

Rassat, A., Land, K., Lahav, O. & Abdalla, F. B., 2007, *MNRAS*, **377**, 1085

Refregier, A., 2003, *Ann. Rev. Astron. Astrophys.*, **41**, 645

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B. & Tonry, J., 1998, *AJ*, **116**, 1009

Riess, A. G., Nugent, P. E., Gilliland, R. L., Schmidt, B. P., Tonry, J., Dickinson, M., Thompson, R. I., Budavári, T. ., Casertano, S., Evans, A. S., Filippenko, A. V., Livio, M., Sanders, D. B., Shapley, A. E., Spinrad, H., Steidel, C. C., Stern, D., Surace, J. & Veilleux, S., 2001, *ApJ*, **560**, 49

Rubin, V. C., Waterman, A. H. & Kenney, J. D. P., 1999, *AJ*, **118**, 236

Sackett, C. A., Kielpinski, D., King, B. E., Langer, C., Meyer, V., Myatt, C. J., Rowe, M., Turchette, Q. A., Itano, W. M., Wineland, D. J. & Monroe, C., 2000, *Nature*, **404**, 256

Schäfer, B. M., 2008, *International Journal of Modern Physics D*, **18**, 173

Scherrer, R. J. & Bertschinger, E., 1991, *ApJ*, **381**, 349

Schlamminger, S., Choi, K., Wagner, T. A., Gundlach, J. H. & Adelberger, E. G., 2008, *Physical Review Letters*, **100**(4), 041101

Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., Schommer, R. A., Clocchiatti, A., Kirshner, R. P., Garnavich, P., Challis, P., Leibundgut, B., Spyromilio, J., Riess, A. G., Filippenko, A. V., Hamuy, M., Smith, R. C., Hogan, C., Stubbs, C., Diercks, A., Reiss, D., Gilliland, R., Tonry, J., Maza, J., Dressler, A., Walsh, J. & Ciardullo, R., 1998, *preprint (astro-ph/9805200)*

Schneider, M. D. & Bridle, S., 2010*a*, *MNRAS*

Schneider, M. D. & Bridle, S., 2010*b*, *MNRAS*, **402**, 2127

Schneider, P. & Weiss, A., 1986, *A&A*, **164**, 237

Schrabback, T. *et al.*, 2009, *ArXiv e-prints*

Scoccimarro, R., Sheth, R. K., Hui, L. & Jain, B., 2001, *ApJ*, **546**, 20

Seo, H.-J. & Eisenstein, D. J., 2003, *ApJ*, **598**, 720

Sersic, J. L., 1968, *Atlas de galaxias australes* (Cordoba, Argentina: Observatorio Astronomico, 1968)

Sheldon, E. S., Johnston, D. E., Scranton, R., Koester, B. P., McKay, T. A., Oyaizu, H., Cunha, C., Lima, M., Lin, H., Frieman, J. A., Wechsler, R. H., Annis, J., Mandelbaum, R., Bahcall, N. A. & Fukugita, M., 2009, *ApJ*, **703**, 2217

Sheth, R. K. & Tormen, G., 1999, *Mon. Not. Roy. Astron. Soc.*, **308**, 119

Simpson, F. & Bridle, S., 2005, *PRD*, **71**(8), 083501

Simpson, F. & Bridle, S., 2006, *PRD*, **73**(8), 083001

Smail, I. & Dickinson, M., 1995, *ApJL*, **455**, L99

Smail, I., Ellis, R. S. & Fitchett, M. J., 1994, *MNRAS*, **270**, 245

Smith, R. E., Peacock, J. A., Jenkins, A., White, S. D. M., Frenk, C. S., Pearce, F. R., Thomas, P. A., Efstathiou, G. & Couchman, H. M. P., 2003, *MNRAS*, **341**, 1311

Smoot & et al., 1992, *ApJ 396, L1*

Song, Y.-S., Zhao, G.-B., Bacon, D., Koyama, K., Nichol, R. C. & Pogosian, L., 2010, *ArXiv e-prints*

Spergel, D. N. *et al.*, 2007, *ApJS*, **170**, 377

Takada, M. & Jain, B., 2004, *MNRAS*, **348**, 897

Takada, M. & White, M., 2004, *ApJL*, **601**, L1

Tereno, I., Semboloni, E. & Schrabback, T., 2011, *A&A*, **530**, A68+

Thomas, S. A., Abdalla, F. B. & Lahav, O., 2010, *Physical Review Letters*, **105**(3), 031301

Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., Krzeminski, W., Mateo, M., Preston, G. W. & Paczynski, B., 1993, *Acta Astronomica*, **43**, 289

van Waerbeke, L. *et al.*, 2000, *Astron. Astrophys.*, **358**, 30

Van Waerbeke, L., Mellier, Y., Erben, T., Cuillandre, J. C., Bernardeau, F., Maoli, R., Bertin, E., Mc Cracken, H. J., Le Fèvre, O., Fort, B., Dantel-Fort, M., Jain, B. & Schneider, P., 2000, *A&A*, **358**, 30

Van Waerbeke, L., White, M., Hoekstra, H. & Heymans, C., 2006, *Astropart. Phys.*, **26**, 91

Velander, M., Kuijken, K. & Schrabback, T., 2011, *MNRAS*, **412**, 2665

Voigt, L. M. & Bridle, S. L., 2010*a*, *MNRAS*, **404**, 458

Voigt, L. M. & Bridle, S. L., 2010*b*, *MNRAS*, **404**, 458

Walsh, D., Carswell, R. F. & Weymann, R. J., 1979, *Nature*, **279**, 381

Weiner, B. J., Phillips, A. C., Faber, S. M., Willmer, C. N. A., Vogt, N. P., Simard, L., Gebhardt, K., Im, M., Koo, D. C., Sarajedini, V. L., Wu, K. L., Forbes, D. A., Gronwall, C., Groth, E. J., Illingworth, G. D., Kron, R. G., Rhodes, J., Szalay, A. S. & Takamiya, M., 2005, *ApJ*, **620**, 595

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I. & Bernstein, G., 2000, *Nature*, **405**, 143

Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I. & Bernstein, G., 2000, *Nature*, **405**, 143

York, D. G. *et al.*, 2000, *AJ*, **120**, 1579

Zhang, P., 2008, *ArXiv e-prints*

Zhang, P., 2010, *MNRAS*, **406**, L95

Zhang, P., Liguori, M., Bean, R. & Dodelson, S., 2007, *Physical Review Letters*, **99**(14), 141302

Zhang, P., Pen, U. & Bernstein, G., 2010, *MNRAS*, **405**, 359

Zhao, G., Pogosian, L., Silvestri, A. & Zylberberg, J., 2009, *Physical Review Letters*, **103**(24), 241301

Zwicky, F., 1933, *Helvetica Physica Acta*, **110**, 6

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, Dr. Sarah Bridle, who has made this work exciting, enriching and never a chore. Your constant support gave me the confidence that a solution would always be found no matter what obstacles stood in the way. Ever generous with time, advice and Matlab hints, you specialise in the five minute chat that makes a month's reading into perfect sense. I am grateful for the opportunities to present my work, speak at conferences and meet other cosmologists that you have encouraged me to take. Thank you Sarah for an enlightening three years.

Thank you to Ofer Lahav, a wise and inspiring second supervisor, always with an eye on the big picture. Thank you to Filipe Abdalla and Hiranya Peiris for the help, encouragement, ideas and inspiration you have provided.

Thanks to all the Group A cosmologists Lisa, Ole, Shaun, Adam & Sam in particular for puzzling over some things with me and explaining many more. Thanks to the whole of Group A for a great three years and thank you to all the past and present occupants of G14 for making UCL such a pleasant place to work day by day - you'll get that sofa one of these days.

I was lucky to get a great set of teachers at school but I'd particularly like to remember Eamon Lewis and Noel Savage. Both kind and inspiring teachers as well as true gents.

It's my friends that have made London life for me, especially Vinnie, Criostoir and Tom for putting up with me at such close proximity.

The epigraph is for Auntie Una & Fr. Sean. "What is the stars?" will always mean more than astronomy.

Mum and Dad, this thesis is dedicated to you. No one could dream of better parents or better company. Even when we can't chat in person over a bottle of wine, I know you're at the end of the phone with news and stories and encouragement. I love you a million zillions.

Finally, thank you Sofie. You more than anyone have made the last three years what they have been for me. Thank you for being with me during the lows and making most of the highs. Thank you for sharing your life with me, for the chats, the political arguments, for all the laughs and for listening to me rant about code. I love you.